Journal of Archaeological Science 53 (2015) 133–147

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Through a filter, darkly: population size estimation, systematic error, and random error in radiocarbon-supported demographic temporal frequency analysis

William A. Brown [a, b, *]

[a] Department of Anthropology, University of Washington, Box 353100, Seattle, WA 98195-3100, USA
[b] The Center for Social Science Computation and Research (CSSCR), University of Washington, Box 353345, Seattle, WA 98195-3345, USA

ARTICLE INFO

Article history:
Received 24 August 2014
Received in revised form
14 October 2014
Accepted 16 October 2014
Available online 23 October 2014

Keywords:
Archaeological demography
Paleontology
Temporal frequency analysis
Radiocarbon age estimation
Monte Carlo simulation

ABSTRACT

Archaeologists are increasingly concerned that the non-linear relationship between the calendric and radiocarbon timelines may introduce anomalous structures into radiocarbon-supported temporal frequency distributions (*tfds*) – time series data describing temporal fluctuations in the frequency of archaeological, paleontological, or other geological deposits. This concern emphasizes a need for improved middle range theory on *tfd* formation, addressing the interaction between several stochastic processes. This paper outlines a Monte Carlo simulation designed to explore the influence of several variables on *tfd* morphology, including the nonlinear calendric-to-radiocarbon age relationship. The results indicate that this non-linear relationship entails greater variance between identically generated *tfds* over some temporal intervals than others but does not predictably lead to *tfd* peaks over these intervals as previously suggested. Additional variance between identically generated *tfds* results from small sample sizes and high values in the underlying *TFD*. Smoothing the *tfd* is a solution not only to calibration curve interference but also to sample size-dependent sampling error.

1. Introduction

In the wake of Paul Ehrlich's *The Population Bomb* (1968), a dystopian treatise on human overpopulation, demographers and population ecologists have paid sustained attention to the population growth dynamics of our species, as well as their implications for past, present, and future well-being (Bloom, 2011; Coale, 1974; Cohen, 1995; Ellison et al., 2012: 757–759; Gerland et al., 2014; Hirschman, 2005; Lee and Tuljapurkar, 2008; Lee et al., 2009; Puleston and Tuljapurkar, 2008; Rogers, 1992). Both population size and growth factor among the chief concerns of demographic and population ecological inquiry (Chamberlain, 2006: 2, 2009: 275–276; Hassan, 1981: 1–2; Micklin and Poston, 2005: 2; Newell, 1988: 3; Preston et al., 2001: 1–16; Skalski et al., 2005: 11), and consequently foundational texts in these fields devote considerable attention to the discussion of methods for describing and estimating population size and growth (Newell, 1988: 8–9; Preston et al., 2001: 1–16; Skalski et al., 2005: 11–15, 289–296); for formalizing the connections between population growth and other topics of demographic concern (e.g., population composition and the age pattern of vital events; Preston et al., 2001: 138–190; Skalski et al., 2005: 26–37); and for fitting population growth models to empirical data and projecting such models into hypothetical futures (Preston et al., 2001: 117–137; Skalski et al., 2005: 15–26, 297–322).

Our ability to identify changes in the size of past populations is relevant to a diverse range of research topics in archaeological and paleontological research, including the identification of colonization and extirpation processes, factors promoting or undercutting the ecological foothold of species, and causality in sociopolitical and economic change and in cultural diffusion and transmission. Over the last four decades, this need has increasingly been met by using temporal frequency distributions (*tfds*) – time series data describing temporal variation in the abundance of archaeological or fossil deposits (Fig. 1) – as proxy census records, or at minimum as records of occupation intensity or presence/absence (see Chamberlain, 2006: 131–132, 2009: 280). Most regions of the world have been subjected to such

Abbreviations: dTFA, demographic temporal frequency analysis; LLN, Law of Large Numbers; MC, Monte Carlo; MRT, middle range theory; *tfd*, a sample temporal frequency distribution; *TFD*, a theoretical temporal frequency distribution, expressed as a probability distribution, from which *tfds* may be sampled; *spd*, summed probability distribution.

* Department of Anthropology, University of Washington, Box 353100, Seattle, WA 98195-3100, USA.
E-mail address: brownw@uw.edu.

http://dx.doi.org/10.1016/j.jas.2014.10.013
0305-4403/

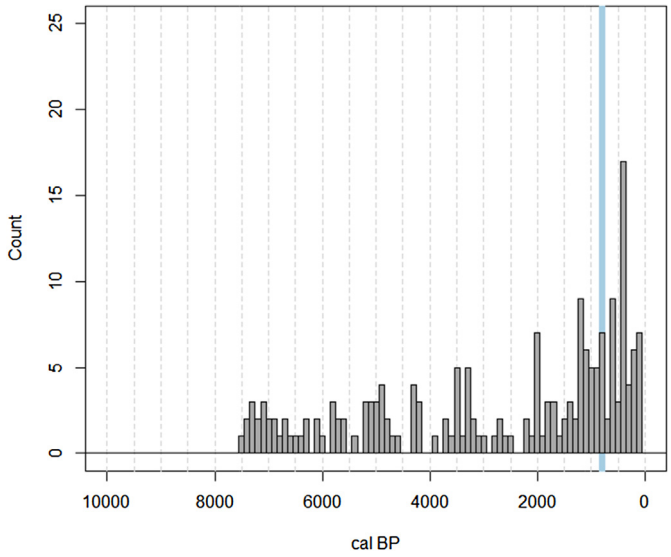

Fig. 1. *tfd* of 187 archaeological site occupations from the Kodiak Archipelago in the Gulf of Alaska, expressed as a histogram (adapted from Maschner et al., 2009b: Fig. 3.5). The blue vertical bar, demarcating the timespan 850–750 cal BP (A.D. 1100–1200), is a century purportedly marked by a dearth of archaeological deposits in Kodiak (Maschner et al., 2009b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

demographic temporal frequency analysis (dTFA), to varying degrees (Table 1).

Research questions addressed in the framework of dTFA have varied widely in temporal scale, focusing on both long-run population processes and relatively brief episodes of demographic change. Among the latter, Maschner and colleagues (2009b: 38–41) have recently applied dTFA to the evaluation of a population replacement model of the Kachemak–Koniag transition in the Kodiak Archipelago, Gulf of Alaska, approximately 900–700 cal BP (Clark, 1998; Fitzhugh, 2003; Maschner et al., 2009b; Mills, 1994; Saltonstall and Steffian, 2006; West, 2011). This cultural change involved the appearance of ceramic technology; a transition from single-to multi-room pit houses; an explosion in typical site size; intensified usage of river corridors and a corresponding increase in the freshwater fishery; changes in symbolic expressions and mortuary practices, likely relating to changes in social and belief systems and ritual practices; and the establishment of institutionalized sociopolitical inequality. A number of archaeologists working in the region have proposed population continuity across this cultural transition and explained it as a process of gradual in situ change (Jordan and Knecht, 1988; Knecht, 1995; Saltonstall and Steffian, 2006), perhaps with exogenous influence from a small number of immigrants or other long-distance contacts (Clark, 1988, 1992, 1998; Fitzhugh, 2003). In contrast, as a component of a synthetic ecological model of cultural and demographic processes spanning the Northeast Pacific Rim, Maschner et al. (2009b) have argued that a dTFA of archaeological site occupations from Kodiak supports Dumond's (1965, 1987, 1988, 1991, 2005) earlier model of cultural change driven by demographic disruption and replacement, in which Kachemak populations largely abandoned the region and were subsequently replaced by Koniag immigrants (see also Hrdlička, 1944). The most striking feature of Maschner and colleagues' dTFA is an alleged paucity of sites during the century 850–750 cal BP, though a closer inspection of their *tfd*s (Fig. 1, adapted from Maschner et al., 2009b: Fig. 3.5.) instead suggests that this paucity came the following century.

Yet, as interest in dTFA grows, archaeologists are also increasingly aware of the approach's limitations and the need to identify

Table 1
Archaeological and paleontological publications featuring temporal frequency distributions.

Area	References
Africa	Ozainne et al., 2014
Near East	Ammerman et al., 1976
Europe	Adams et al., 2001; Armit et al., 2013; Collard et al., 2010b; Crombé and Robinson, in press; Gamble et al., 2004, 2005; Gkiasta et al., 2003; González-Sampériz et al., 2009; Hinz et al., 2012; Housley et al., 1997; Kerr and McCormick, 2014; Liedgren et al., 2007; Martínez et al., 1997; Mellars and French, 2010; Naudinot et al., in press; Riede, 2008, 2009; Schmidt et al., 2012; Shennan and Edinborough, 2007; Shennan et al., 2013; Tallavaara and Seppä, 2012; Tallavaara et al., 2010; Timpson et al., in press; Turney et al., 2006; Whitehouse et al., 2014; Wicks and Mithen, 2014; Wicks et al., 2014; Woodbridge et al., 2014
Eastern Asia	Barton et al., 2007; Brantingham et al., 2004; Crema, 2012; Fiedel and Kuzmin, 2007; Kim and Bae, 2010; Kuzmin and Keates, 2005; MacPhee et al., 2002; Rhode et al., in press; Surovell et al., 2005; Wang et al., 2014
North America	Adams et al., 2001; Albanese and Frison, 1995; Anderson and Freeburg, 2013; Ballenger and Mabry, 2011; Batt and Pollard, 1996; Bever, 2006; Buchanan et al., 2008, 2011; Collard et al., 2008, 2010a; Erlandson et al., 1992, 2001; Esdale, 2008; Fitzhugh, 2003; Grier, 2006; Guthrie, 2006; Haggarty et al., 1991; Haynes, 1969; Hutchison and McMillan, 1997; Kelly et al., 2013; Lepofsky et al., 2005; Louderback et al., 2011; Maschner et al., 2009a, 2009b; Mason et al., 2001; Mullen, 2012; Munoz et al., 2010; Peros et al., 2010; Potter, 2008; Rasic and Matheus, 2007; Steele, 2010; Story and Valastro, 1977; Surovell et al., 2009; Taylor et al., 2011; Weninger et al., 2011
South America	Delgado et al., in press; Gil et al., 2005; Rick, 1987; Moreno et al., 2009; Williams et al., 2008
Oceania	Adams et al., 2001; Holdaway and Porch, 1995; Holdaway et al., 2008; Kirch, 1985; Kirch et al., 2012; McFadgen et al., 1994; Marwick, 2009; Miller et al., 1999; Mulrooney, 2013; Smith and Ross, 2008; Smith et al., 2008; Turney and Hobbs, 2006; Williams, 2012; Williams et al., 2008, 2010, 2013

the conditions and constraints on its valid pursuit. In response, researchers now strive for a more explicit, comprehensive, and sufficiently compelling body of middle range theory (MRT) to support and regulate dTFA, one that outlines the system of transformations through which paleopopulation size curves pass in the process of becoming archaeological or paleontological *tfd*s. This effort, outlined in Section 1.1, has been underway since Rick's (1987) article on the population dynamics of pre-ceramic Peru and will likely continue for some time. This paper focuses on one aspect of this work, the unique challenges that confront dTFA supported by radiocarbon (^{14}C) age estimation (Section 1.2).

1.1. The MRT of dTFA

The validity of dTFA rests on the proposition that larger populations tend to discard a greater abundance of materials than do smaller ones (Rick, 1987). By implication, *tfd*s generated from samples of time-stamped archaeological or paleontological deposits should track temporal variation in past population size, all else being equal.

Of course, all else is rarely equal. Rick (1987) qualified this strong assumption almost as soon as he articulated it, identifying a number of potential confounding factors each warranting dedicated efforts at mitigation or control. In the two and a half decades following Rick's article, disputed interpretations have become a standard in

dTFA, one of the most recent and heated arising in response to Buchanan and colleagues' (2008) rejection of a Paleoindian population response to the hypothetical Younger Dryas (YD) extraterrestrial impact event (see Anderson et al., 2008; Bamforth and Grund, 2012; Buchanan et al., 2011; Collard et al., 2008; Culleton, 2008; Firestone et al., 2007; Jones, 2008; Kennett and West, 2008; Kennett et al., 2008; Steele, 2010; Weninger et al., 2011).

General factors confounding dTFA can be divided between those that introduce systematic error into the population size estimate and those that introduce random error (Table 2).

1.2. Systematic and random error and non-demographic structures in ^{14}C-supported tfds

Special issues in dTFA arise when archaeologists or paleontologists depend on ^{14}C age-stamped samples for their analysis. In this case, two additional factors become relevant: secular variation in the concentration of atmospheric ^{14}C (Sonett and Finney, 1990), and instrument measurement error (Walker, 2005), each entailing unique computational demands for *tfd* construction and analysis. These demands have increasingly been met by ^{14}C age calibration and the construction of summed probability distributions (*spd*s).[1] With this type of *tfd*, the probability density characterizing each probabilistic age estimate in the sample is summed at each location along the timeline:

$$u(t) = \sum_{i=1}^{n} f_i(t) \quad (1)$$

where $u(t)$ is the sum of all units' probabilistic age estimates at time t and $f_i(t)$ is the probability density for the i th data point in the sample at time t. In effect, *spd*s provide a best guess regarding the temporal frequency of a sample of units, given the uncertainty that characterizes ^{14}C-based age estimation.

Fig. 2 presents a *spd* comprising 166 site occupations from Kodiak. Data were compiled by the present author, updating previous databases compiled by Haggarty, Erlandson, and colleagues (Erlandson et al., 1992; Haggarty et al., 1991), Mills (1994), and Fitzhugh (2003). Dates excluded from analysis include strictly geologic dates; modern dates and dates reported without lab error; dates lacking sufficient contextual information to evaluate their archaeological validity; dates widely dismissed by archaeologists working in the region; and dates on marine fauna or suspected marine fauna. For each site in the sample, clusters of insignificantly different age estimates were pooled following the method proposed by Ward and Wilson (1978), to minimize the overrepresentation of intensively dated sites. It is worth noting that Maschner and colleagues' *tfd* (Fig. 1) resembles the *spd* shown in Fig. 2 in broad outline, as well as in several details, and particularly the brief paucity of sites 750–650 cal BP.

While ^{14}C age calibration and the use of *spd*s are regarded as best practices in ^{14}C-supported dTFA, the *spd*s they generate unfortunately exhibit anomalous peak-and-trough structures arising from several recognized sources of error:

1. *Sampling error. tfd*s, including *spd*s, are sample distributions and as such approximate the shape of their underlying probability

Table 2
A typology of confounding factors in dTFA.

Rick's (1987) typology of error in dTFA	Definition	References
Creation	Systematic error: diachronic change in the prevailing per capita deposition rate Random error: stochastic variation around the deposition rate (a Poisson process)	Peros et al., 2010; Rick, 1987; Taylor et al., 2011
Preservation	Systematic error: (1) time transgressive site destruction; (2) diachronic change in the prevailing rate of site destruction Random error: stochastic variation around the probability of deposit survival	Albanese and Frison, 1995; Ballenger and Mabry, 2011; Clevis et al., 2006; Mandel, 1995; Michczyńska et al., 2003; Rick, 1987; Surovell and Brantingham, 2007; Surovell et al., 2009; Waters and Kuehn, 1996
Investigation	Systematic error: over- or underrepresentation of archaeological deposits from certain periods of time resulting from (a) the use of particular survey methods, or (b) heavy research focus on particular periods of special interest Random error: (1) Random sampling error; (2) failure to include accurate dates or exclude spurious ones resulting from the implementation of imperfect sample hygiene protocols	Ballenger and Mabry, 2011; Batt and Pollard, 1996; Chamberlain, 2006, 2009; Dean, 1978; Faught, 2008; Haynes et al., 2007; Kennett et al., 2008; Michczyńska et al., 2003; Pettitt et al., 2003; Rick, 1987; Rieth and Hunt, 2008; Schiffer, 1987; Shott, 1992; Surovell and Brantingham, 2007; Waters and Stafford, 2007; Williams, 2012

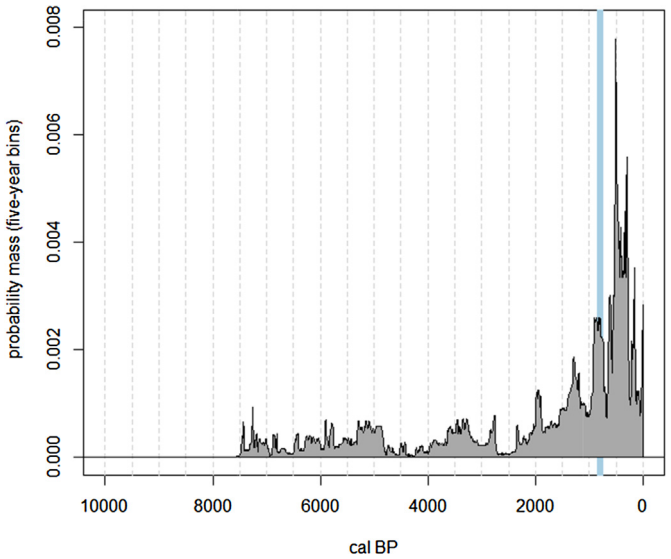

Fig. 2. *spd* of 166 site occupations from the Kodiak Archipelago. Ages were calibrated using the IntCal13 (Reimer et al., 2013) calibration curve and script written by the author in the R programming language. See Supplemental Material for data. The blue vertical bar is as in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

[1] In the dTFA literature, the term 'cumulative probability distribution/function' has occasionally been used in place of 'summed probability distribution/function.' This usage is unfavorable, having the potential to be confused with either of two standard statistical terms – '(cumulative) density function' (df or cdf, expressed $F(x)$ in mathematical notation) and 'cumulative probability function' – referring to the probability $Pr(X \leq x)$ for continuous and discrete variates, respectively.

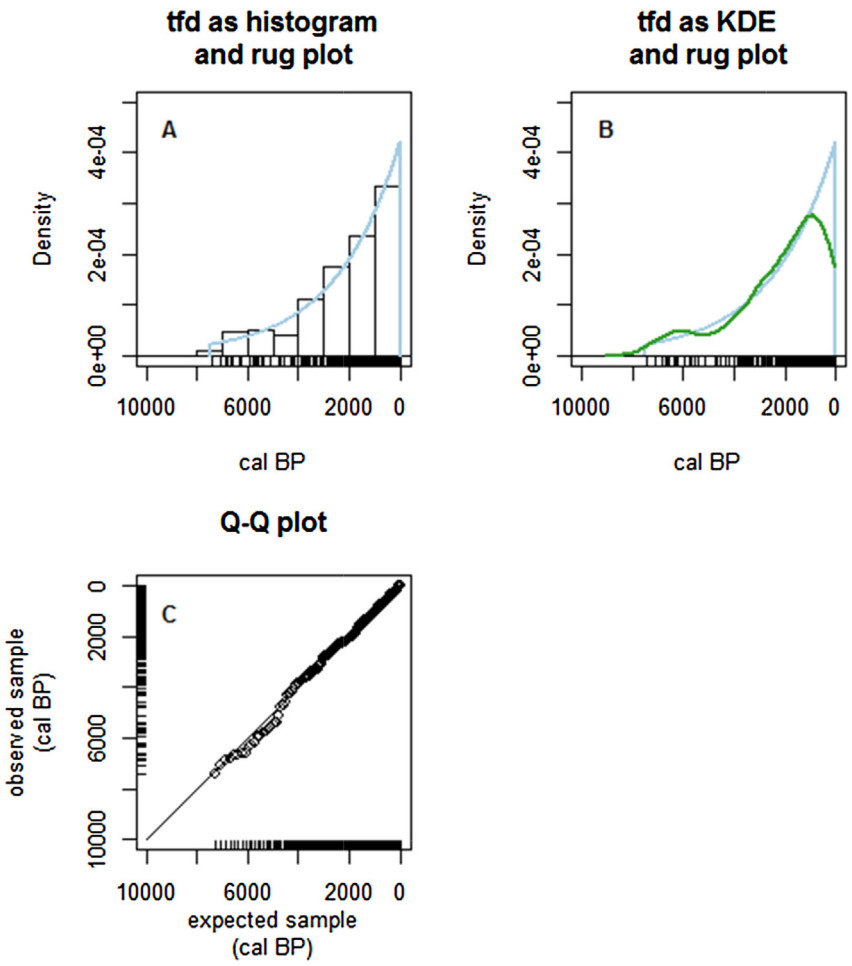

Fig. 3. *A*: *tfd*, expressed as a histogram and rug plot, describing a sample of 200 observations randomly sampled from a truncated exponential *TFD* (light blue line). *B*: *tfd*, expressed as a KDE (green line) and rug plot, describing the same sample of 200 observations as in Fig. 3:A. *C*: Quantile–quantile plot comparing the observed sample from Fig. 3:A and 3:B to the expected sample. The diagonal line describes an ideal 1:1 relationship. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distribution (*TFD*) imperfectly. Fig. 3 illustrates the morphological disparities standing between a truncated exponential *TFD* and a *tfd* of 200 observations randomly drawn from it. Anomalous gaps and concentrations between observations are evident in the rug plot expression of the *tfd*, as indicated in the quantile–quantile plot (Fig. 3:C), informing the artificial peaks, valleys, and plateaus observed in both the histogram (Fig. 3:A) and kernel density estimate (KDE; Fig. 3:B). The influence of sampling error on *tfd* shape has been explored by Michczyńska and colleagues (Michczyńska and Pazdur, 2004; Michczyńska et al., 2003), Williams (2012), Armit et al. (2013), Rhode and colleagues (in press), and Contreras and Meadows (in press) and is also apparent in Bartlein and colleagues' (1995: Fig. 2) uncalibrated *tfd* simulations.

2. *Dispersive calendric-to-^{14}C age mapping.* Each calendric age corresponds with a range of possible ^{14}C ages rather than a single point. For example, the calendric age 12,000 cal BP corresponds with a range of ^{14}C ages specified as *N*(10,237, 47) by IntCal13 (Reimer et al., 2013). Figs. 4:A and 5:A illustrate the disparities that arise between ^{14}C ages having an identical calendric age of 12,000 cal BP.
3. *Instrument measurement error.* As is true in the measurement sciences in general, the instruments employed in ^{14}C age measurement entail random error that obscures the true ^{14}C age of dated specimens. As a result, even repeated measurements on a single specimen may produce disparate estimates. This instrument measurement error interacts with dispersive calendric-to-^{14}C age mapping to further extend the range of ^{14}C ages corresponding with any given calendric age, leading to a propagation of error. In Fig. 4:A, the true ^{14}C ages of all three simulated observations have each been offset by a value following a normal distribution specified as *N*(0, 50). Fig. 5:B shows the distribution of 10,000 measured ^{14}C ages corresponding with the true ^{14}C ages shown in Fig. 5:A, robustly illustrating the pattern that emerges from this propagation of error. Importantly, when the three age estimates shown in Fig. 4:A are calibrated (Fig. 4:B), it becomes clear that the probabilistic calendric age estimates from which *spd*s are generated can be profoundly dislocated by this propagation of error. Bamforth and Grund (2012), Armit et al. (2013), Kerr and McCormick (2014), Wicks and Mithen (2014), and Contreras and Meadows (in press) have incorporated this propagation of error into their simulation studies using OxCal's `R_Simulate` function (Bronk Ramsey, 2009a,b). In fact, this effect accounts entirely for the differences observed between two identically generated *spd*s in Bamforth and Grund's (2012: Fig. 1) study in the absence of sampling error. McFadgen et al. (1994) have also explored the influence of instrument measurement error on *tfd* morphology.

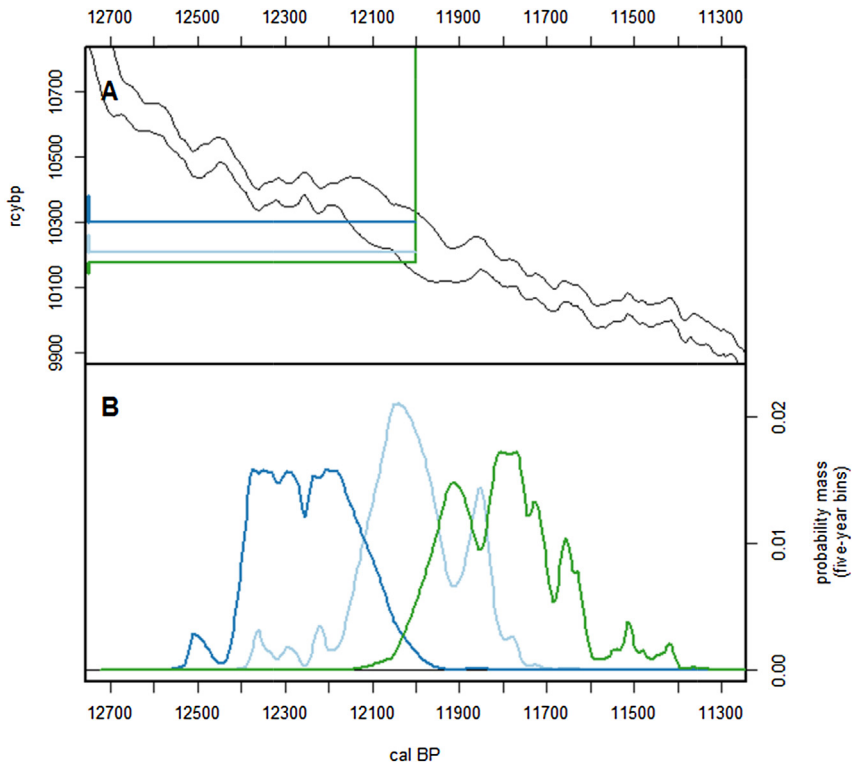

Fig. 4. *A*: Dispersive calendric-to-^{14}C age mapping for three simulated observations having an identical true calendric age of 12,000 cal BP, showing additional dispersion through instrument measurement error. Dispersive mapping for the true age 12,000 cal BP follows the distribution *rcyBP*~$N(\mu = 10{,}237, \sigma = 47)$ per IntCal13 (Reimer et al., 2013), with an instrument measurement offset following a normal distribution specified as $N(\mu = 0, \sigma = 50)$. *B*: calibrated probability distributions for the three ^{14}C age estimates presented in Fig. 4:A, each assuming a standard error of 50 ^{14}C years.

4. *Calibration interference.* Calibration as it is currently implemented involves mapping the entire probability distribution associated with a ^{14}C age estimate through the calibration curve into calendric time. In the process, the shape, location, and scale of the distribution are transformed. Because the shape of the calibration curve varies over time, the degree to which these three distributional properties are transformed likewise varies along the timeline (Guilderson et al., 2005; Pazdur and Michczyńska, 1989; Steier et al., 2001; Stuiver and Reimer, 1989, 1993; Weninger, 1986; Weninger et al., 2011). Fig. 6 illustrates this time-variant effect. While the shape and scale of two probability distributions is identical along the ^{14}C timeline, their respective expressions along the calendric timeline is notably disparate: the younger calendric estimate is more precise than its corresponding, uncalibrated counterpart and is also asymmetrical, whereas the older calendric estimate is not only more diffuse than its uncalibrated counterpart but also shows multiple, asymmetrically distributed modes. The influence of this error on the shape of ^{14}C-supported *tfd*s has been explored by McFadgen et al. (1994), Michczyńska and colleagues (Michczyńska and Pazdur, 2004; Michczyńska et al., 2003), Williams (2012), Bamforth and Grund (2012), Armit et al. (2013), Kerr and McCormick (2014), Wicks and Mithen (2014), and Contreras and Meadows (in press).

In effect, these four sources of error entail a potential to introduce structural equifinalities into the shape of the *tfd*, with some structures representing real demographic processes and others false ones. At the same time, the appearance of real demographic structures may be dampened by both systematic and random error.

In response to this potential for error, a handful of simulation experiments have been conducted to explore the influence of one (Bartlein et al., 1995; Rhode et al., in press), two (McFadgen et al., 1994; Michczyńska and Pazdur, 2004; Williams, 2012), or three (Bamforth and Grund, 2012; Kerr and McCormick, 2014; Wicks and Mithen, 2014) of these sources of error. However, exploring the interaction of all four factors has only recently begun (Armit et al., 2013; Contreras and Meadows, in press). In addition, the robustness of these experiments has for the most part been limited by unfavorably small simulation sizes, involving either the characterization of a single simulation run (Armit et al., 2013; Bamforth and Grund, 2012; Bartlein et al., 1995; Contreras and Meadows, in press; Kerr and McCormick, 2014; McFadgen et al., 1994; Wicks and Mithen, 2014) or comparisons of between two and one hundred runs (two: Bamforth and Grund, 2012; five: Contreras and Meadows, in press; ten: Armit et al., 2013; thirty: Williams, 2012; one hundred: Rhode et al., in press). Michczyńska and colleagues' Monte Carlo (MC) simulations, based on 10,000 runs, are the noteworthy exceptions (Michczyńska and Pazdur, 2004; Michczyńska et al., 2003).

The next section presents a general framework for a flexible MC simulation designed to address these critical gaps. It begins by considering the motivation for and properties of effective MC simulations. Section 3 presents the results of two permutations of this approach, highlighting the influence of size-dependent sampling error and *TFD* shape on inter-run *tfd* variability.

2. A MC simulation approach

2.1. Motivation, goals, and ideal properties of MC simulations

Frequently, modeling complex systems of causal relationships eludes closed-form expression, inhibiting the possibility of predicting their outcomes through analytical solution (Roberts and

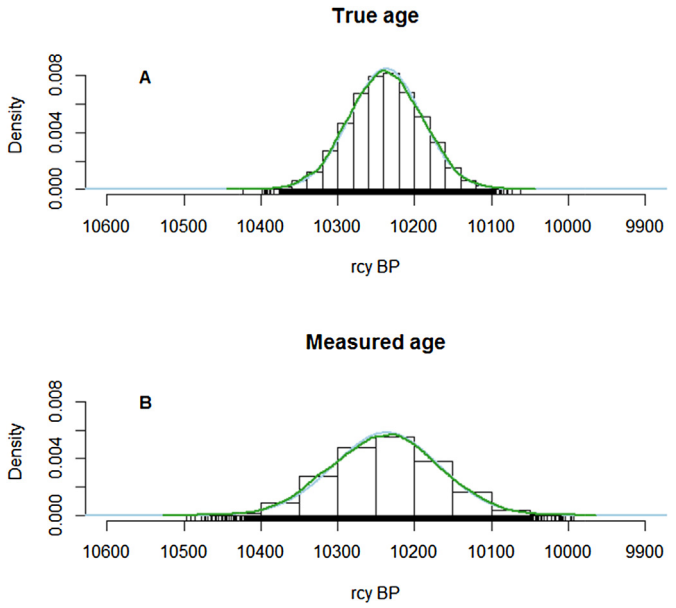

Fig. 5. *A*: Distribution of true ^{14}C ages for 10,000 simulated observations having an identical true calendric age of 12,000 cal BP. These observations follow the normal model specified as $rcyBP_{true} \sim N(\mu = 10{,}237, \sigma = 47)$ (light blue line), based on IntCal13 (Reimer et al., 2013). The sample mean and standard deviation are 10,237.42 and 47.64, respectively. A KDE of this sample (green line) closely approximates the model distribution. *B*: Distribution of measured ^{14}C ages for the same 10,000 observations as in Fig. 5:A. Simulated measurement error follows the normal model specified as $offset \sim N(\mu = 0, \sigma = 50)$. Propagation of error implies a normal model specified as $rcyBP_{measured} \sim N(\mu = 10237, \sigma = \sqrt{47^2 + 50^2} = 68.62)$ (light blue line). The sample mean and standard deviation are 10,238.27 and 69.75, respectively. A KDE of the sample (green line) once again closely approximates the model distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Casella, 2004; Rubinstein and Kroese, 2008; Thomopoulos, 2013). This is true, for example, when the output follows an unknown probability distribution that combines the effects of separate but interacting stochastic processes, each governed by its own data generating process (DGP). In the context of ^{14}C-supported dTFA, exploring the combined influence of three stochastic processes — random sampling, calendric-to-^{14}C-age mapping, and instrument measurement error — on the shape of *spds* provides a case in point. This compound DGP, in combination with the complicated transformation entailed by calibration, render any effort to infer the *spd* expression of a known *TFD* analytically intractable.

When analytical solutions cannot be obtained for problems such as this, MC simulation offers a means of approximating the solution within a reasonable degree of uncertainty. Customarily, MC simulations conform to the following, general procedure:

1. Design an algorithm that
   a. automates the passage of output from earlier components of the system to later ones;
   b. dictates realistic functional relationships, whether deterministic or probabilistic, between each component's input and output variables;
2. Hold certain variables in the model system constant while allowing others to vary within meaningful limits;
3. Initialize the simulation by generating a large set of random observations and passing them through the algorithm, recording the final output for each pass (as well as intermediate outcomes, if of interest);
4. Calculate one or more summary statistics for the set of outcomes and treat these as approximations for the elusive analytical solution. Such statistics are labeled 'plug-in estimators,' to distinguish them from both their analytic counterparts and ordinary sample statistics. Plug-in estimators may include single-point values (e.g., mean, median, variance, percentiles) or entire sample distributions (e.g., histograms, KDEs).

With large enough sample sizes, MC plug-in estimators take advantage of the Law of Large Numbers (LLN), which states that sample statistics should converge toward their underlying parameters as sample size increases (Christian and Casella, 2010). In practice, the efficiency with which computers can generate random numbers allows for the implementation of incredibly large simulations and consequently very precise plug-in estimators.

However, when sample size is itself a parameter of interest in the model system, as it is when sampling error exerts a major influence on the shape of a sample distribution, variation reduction is no longer the goal of the simulation. Rather, exploring the influence of sample size on variation between samples becomes an analytical target. Even in this case, however, achieving a robust and precise estimate of variation between samples requires the simulation of a large set of samples, warranting a two-step elaboration on the basic MC framework:

1. Generate a single run of random inputs, manipulating sample size rather than favoring a large one, and calculate the plug-in estimator(s) of interest from its output;
2. Repeat the first step a large number of times, collecting plug-in estimator(s) of interest from each run, then calculate second-order, inter-run estimators from the collection of per-run estimators. This is the MC equivalent to solving for the parameters of an unknown sampling distribution.

2.2. Simulating variation between ^{14}C *spds*

Past simulation studies exploring ^{14}C *spds* have emphasized the influence of several variables on *spd* morphology. These include the shape, scale, and location of the underlying *TFD* (assuming for the sake of argument that this is proportional to the population size curve, $N(t)$); sample size in a random sampling framework; the shape of the ^{14}C-to-calendric-age curve (i.e., the calibration curve), used for both calendric-to-^{14}C-age translation and subsequent calibration; and the magnitude of instrument measurement error (Armit et al., 2013; Bamforth and Grund, 2012; Bartlein et al., 1995; Contreras and Meadows, in press; Kerr and McCormick, 2014; McFadgen et al., 1994; Michczyńska and Pazdur, 2004; Rhode et al., in press; Wicks and Mithen, 2014; Williams, 2012).

The main constraint on most of these past simulations' ability to achieve robust estimators of inter-run variance has been a lack of automation, often requiring the researchers themselves to transfer output from one computing environment to another, as well as performing some of the simulation steps by hand. For example, Bamforth and Grund (2012) recorded the true ages of their simulated samples in an unspecified environment, then passed these to OxCal 3.10 and 4.1 to translate them into offset ^{14}C age estimates, and finally passed these to CALIB 5.1 and 6.0 to generate *spds*. Similarly, Williams (2012) resampled from an extensive database of archaeological ^{14}C dates in an unspecified environment and generated one *spd* per run in OxCal 4.1, for 210 separate runs. Contreras and Meadows (in press) digitized realistic *TFD*s with Plot Digitizer, generated random samples of varying size from these using an algorithm in R, and passed these to OxCal 4.2.3 to translate them into offset ^{14}C age estimates and calibrate them. This constraint has been due in large part to archaeologists' dependence on one of the major calibration programs — OxCal, CALIB, or CalPal — to generate ^{14}C *spds*. Fortunately, such dependence is not compulsory, as Shennan

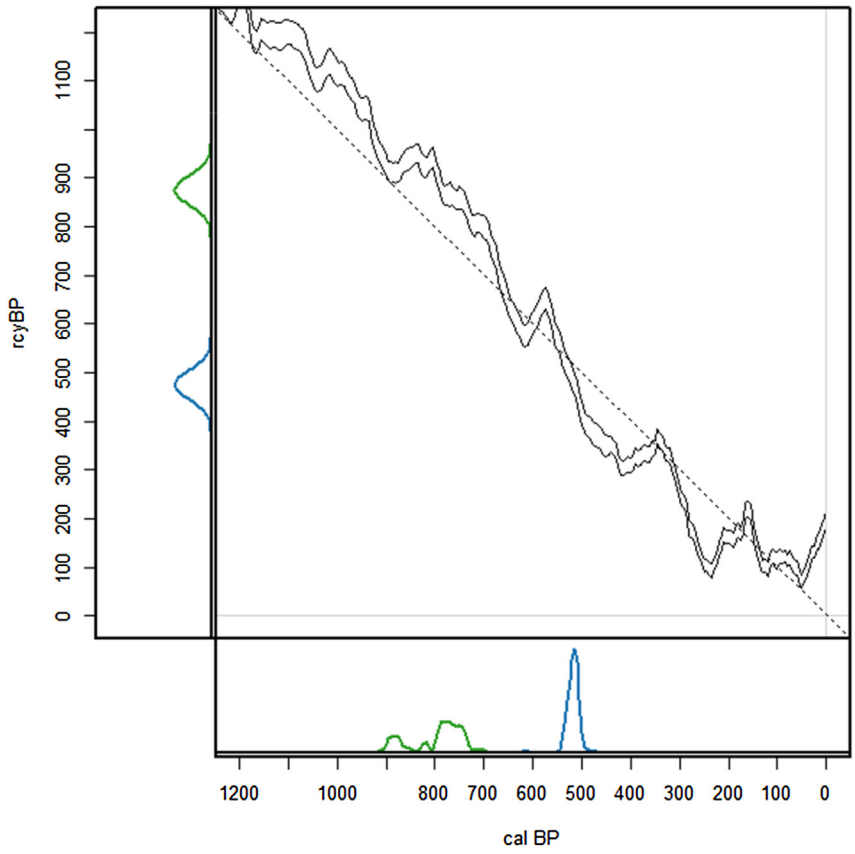

Fig. 6. Two hypothetical ^{14}C age estimates with identical precisions on the ^{14}C age axis (vertical): 475 ± 30 rcyBP (blue curve) and 875 ± 30 rcyBP (green curve). IntCal13 (Reimer et al., 2013) is shown in the main panel (solid black lines), along with an ideal 1:1 relationship between calendric and ^{14}C time (dashed line). The varying slope of the calibration curve mediates the mapping relationship between ^{14}C and calendric age estimates (horizontal axis). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al. (2013) have demonstrated, and the potential power that we might hope to achieve by conducting simulation studies of *spd* formation in a single computing environment is well-illustrated by Michczyńska and colleagues' [2004] unparalleled 10,000-run simulation in GdCALIB. The free availability of the R programming language now puts the potential for fully automated, single-environment simulations at the fingertips of archaeologists. Such well-integrated simulation studies would take the researcher out of the process save for in choosing how to manipulate the parameters of interest. This could result in a vast improvement in simulation efficiency and a significant reduction in user error.

The following presents the general framework of a MC algorithm, implemented in R and designed to be flexible enough to accommodate the experimental manipulation of those parameters listed above. In general, the DGP underlying this algorithm combines the following elements[2]:

- A set of runs is generated, each run comprising a sample of observations randomly drawn from an identical *TFD*. Each observation simulates the true calendric age of a single unit of observation (as in Fig. 3):

$$calBP_{true,\ 1,\ldots,n} \sim TFD(\theta) \tag{2}$$

where the *TFD* is specified as a function of a vector of parameters θ.

- Each true calendric age is mapped into ^{14}C time according to the relationship

$$RCYBP_{true,\ 1,\ldots,n} \sim N(\mu_{CC}(calBP), \sigma_{CC}(calBP)) \tag{3}$$

where $\mu_{CC}(calBP)$ is the mean ^{14}C age corresponding with the true calendric age (equivalent to $r(t_i)$ in Bronk Ramsey, 2009b: Eq. 22) and $\sigma_{CC}(calBP)$ is the accompanying standard deviation (equivalent to $s(t_i)$ in Bronk Ramsey, 2009b: Eq. 22), as dictated by a specified calibration curve.

- True ^{14}C ages are offset to simulate instrument measurement error (*IME*, equivalent to s_i in Bronk Ramsey, 2009b: Eq. (22)) according to the dispersive relationship

$$RCYBP_{measured,\ 1,\ldots,n} \sim N\left(RCYBP_{true,\ 1,\ldots,n}, IME\right) \tag{4}$$

($RCYBP_{measured,i}$ being equivalent to r_i in Bronk Ramsey, 2009b: Eq. (22)). The combined effect of this and the preceding step is a propagation of error as discussed in Section 1.1, such that.

$$RCYBP_{measured,1,\ldots,n} \sim N\left(\mu_{CC}(calBP), \sqrt{\sigma_{CC}(calBP)^2 + IME^2}\right) \tag{5}$$

whose probability density follows the posterior probability function given by Bronk Ramsey (2009b: Eq. 22).

[2] Simulation code in the R programming language is available at https://github.com/brownwuw/radiocarbon_spd_DGP_MC_simulation_code/blob/master/MCsim.functions.

- Measured ^{14}C ages are calibrated according to an algorithm similar to that used by OxCal, CALIB, and most recently Shennan et al. (2013). (For debate regarding the merits of different calibration algorithms, see Buchanan et al., 2011; Culleton, 2008; Steele, 2010; Weninger et al., 2011.) Once calibrated, the probability distributions for all observations in a single run are then summed per Eq. (1), then rescaled so that the area under the *spd* sums to 1.
- Plug-in estimators are calculated summarizing the distributional properties of all runs' *spd* values at five-year intervals along the calendric timeline.

To demonstrate the utility and flexibility of this MC framework, I focus here on an experiment exploring the combined influence of calibration interference, sampling error, and *TFD* shape on *spd* morphology. The eight sample sizes considered (n = 30, 50, 100, 200, 500, 1000, 1500, and 2000) span a range typical for studies in dTFA. The two *TFD*s considered (a uniform and an exponentially increasing *TFD*) are both truncated at 7500 and 0 cal BP to address the range of time covered by Kodiak's known occupation history (see Figs. 1 and 2). These *TFD* morphologies simulate a stationary population and a stable population growing at a constant rate of approximately 0.7064 people per thousand people per year, respectively. As previous researchers have observed (Bartlein et al., 1995; McFadgen et al., 1994; Bamforth and Grund, 2012), the uniform distribution affords an ideal means of demonstrating the intrusion of artificial structures into *tfd*s, given its structureless morphology. A single calibration curve is considered (IntCal13; Reimer et al., 2013). Finally, a single value for instrument measurement error was considered (50 ^{14}C years), approximating the median and mean standard errors for the set of Kodiak dates illustrated in Fig. 2.

For each sample size and *TFD*, 2000 runs were generated per Eq. (2), for a total of 32,000 runs processed. With a maximum sample size of 2000 observations per run and two *TFD*s considered, 8 million individual age estimates were generated in total. Each observation was mapped onto the ^{14}C timeline and offset per Eqs. (3) and (4), then calibrated. All observations per run were summed per Eq. (1) and rescaled so that the area under each *spd* sums to 1, in other words such that each *spd* comprises a time series of sample proportions. The inter-run mean sample proportion and four percentiles (2.5%, 25%, 75%, and 97.5%) were calculated at five-year intervals along the calendric timeline.

3. Results

Figs. 7 and 8 present the simulation output for all sixteen simulations (eight sample sizes per model *TFD*). Four patterns emerge from their examination:

First, the inter-run mean of sample proportions closely approximates the underlying *TFD*, as expected given the LLN and the 2000 runs per simulation. However, two departures from this pattern are worth noting, one at each end of the model *TFD*:

- Lower-than-expected inter-run means are indicated toward the earlier end of each distribution, resulting from the dispersive random error introduced by ^{14}C age translation and instrument offset. In effect, this dispersion has led to the extension of *spd* sample mass beyond the sharply defined limit of the underlying *TFD*.
- A high degree of random error in the inter-run mean of sample proportions, as well as in their percentiles, is indicated over the last 500 years of each distribution. Williams (2012) noted similarly high variance in his bootstrap samples for the same temporal interval. This period of heightened inter-run variability is likely a result of the marked increase in precision of the calibration curve over this interval (Fig. 9), which would act not only to decrease the dispersion evident in the mapping of calendric into ^{14}C ages (Eq. (3)) but also to increase the precision

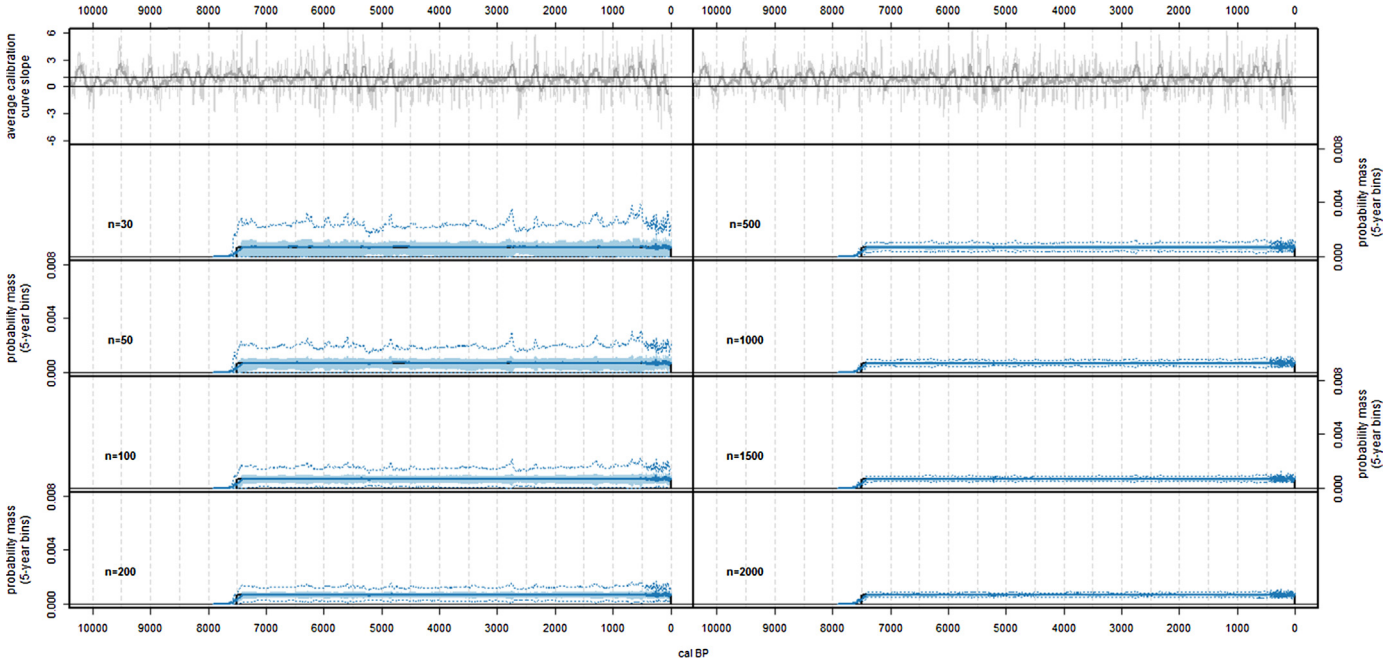

Fig. 7. Results of eight MC simulations, showing inter-run variation between 2000 identically generated *spd*s per sample size, based on a uniform *TFD* (black line). The dotted blue lines show the 2.5% and 97.5% boundaries, while the light blue polygon delineates the interquartile range. The 10- and 100-year moving average slopes of the IntCal13 (Reimer et al., 2013) calibration curve are shown at top (light and dark gray curves, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

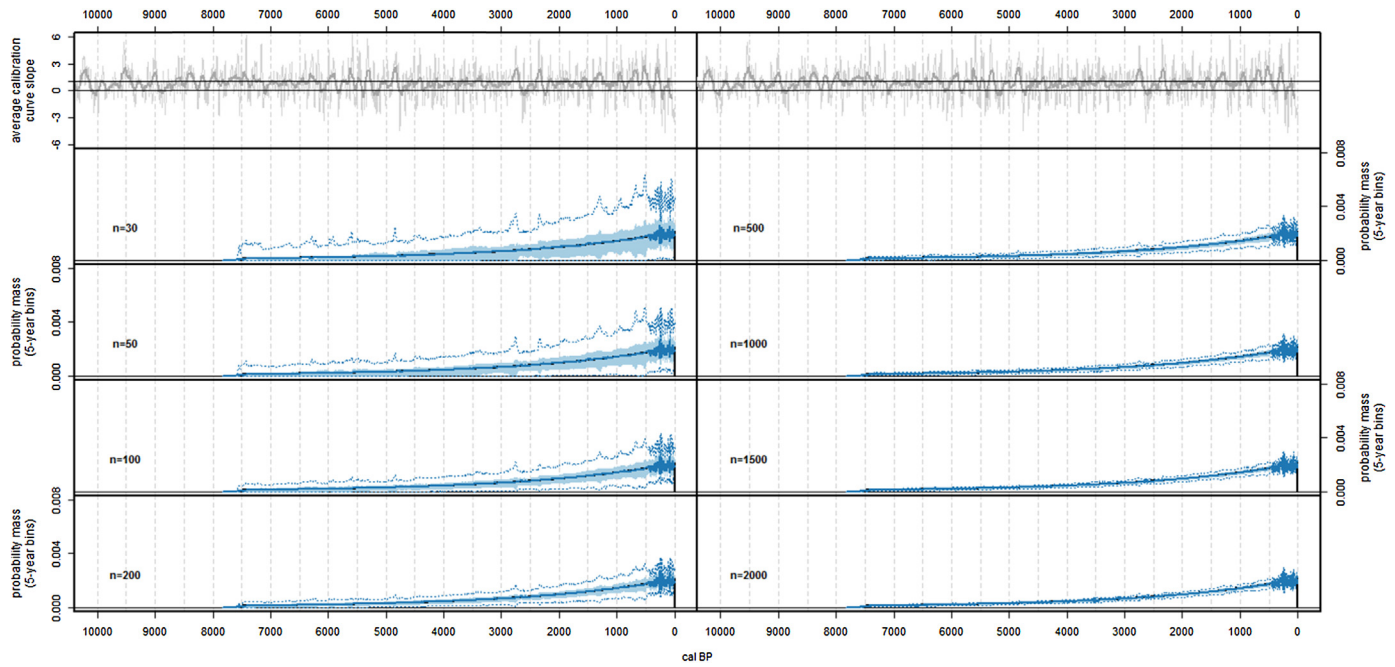

Fig. 8. Results of eight MC simulations, showing inter-run variation between 2000 identically generated *spds* per sample size, based on a truncated exponential *TFD* (blue line). Visual specifications are as in Fig. 7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of calibrated age estimates' probability distributions (as in Fig. 6).

Second, the magnitude of variation between runs decreases as sample size increases, converging toward the underlying *TFD*, as expected given the LLN. However, for small and even moderate sample sizes, the lower percentile boundaries fall at or hover slightly above 0. For the uniform *TFD* simulations, the lower 25% boundary has lifted away only by $n = 100$ and the 2.5% boundary only by $n = 200$. A similar pattern holds for the exponential *TFD* simulations, although the earlier end of the curve remains low in response to the low *TFD* values at this point. Additionally, the inter-run mean falls above the 75% boundary near the early end of the exponential *TFD* simulation at $n = 30$.

Third, in the case of the exponential *TFD* simulations, the dispersion of sample proportions generally increases as the mass of the underlying *TFD* value increases, producing the curved wedge shape characterizing the percentile boundaries shown in Fig. 8. A similar pattern should be expected for any *TFD* exhibiting departures from uniform shape over the timeline. Fig. 10 illustrates *spds* for two simulation runs ($n = 200$) out of the 2000 illustrated in Fig. 8. Here it is worth noting that, while both *spds* are the products of an identical DGP, their inter-run disparities are greatest toward the later, higher-density end of the underlying *TFD*.

Fourth, the fluctuations observed in the percentile estimators over time indicate positional stability across simulations, this pattern being particularly obvious in the uniform *TFD* simulations (Fig. 7). Furthermore, these fluctuations appear to correspond with fluctuations in the average calibration curve slope, as already noted by Michczyńska and Pazdur (2004) and as observed in Fig. 11. On this point, a revision of previous statements about the nature of calibration interference is warranted (see Bamforth and Grund, 2012; McFadgen et al., 1994; Williams, 2012). Temporal intervals characterized by steep calibration curve slopes do not always exhibit anomalously high sample proportions but rather a greater range of variability between *spds*, on both sides of the underlying *TFD* value. In effect, the positional stability of high- and low-variability intervals is an artifact of calibration interference (as in Fig. 6), whereas their direction and magnitude constitute random error stemming from the combined operation of sampling error (as in Fig. 3) and the dispersive transformation of true calendric ages into measured ^{14}C age estimates (as in Fig. 4:A).

4. Discussion

The results of the MC simulations presented here reinforce previous simulation-based impressions (1) that the nonlinear relationship between the calendric and ^{14}C timelines introduces artificial, time-dependent structures into *spds*, and (2) that it is

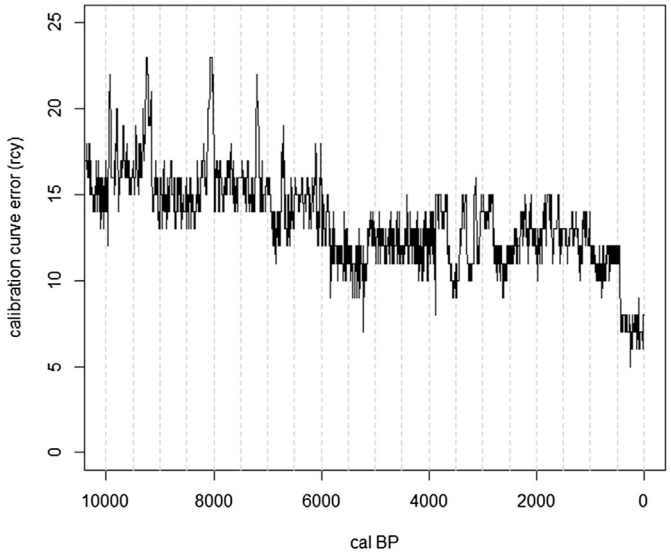

Fig. 9. Standard deviation around the IntCal13 calibration curve (Reimer et al., 2013), over the last 10,000 calendric years.

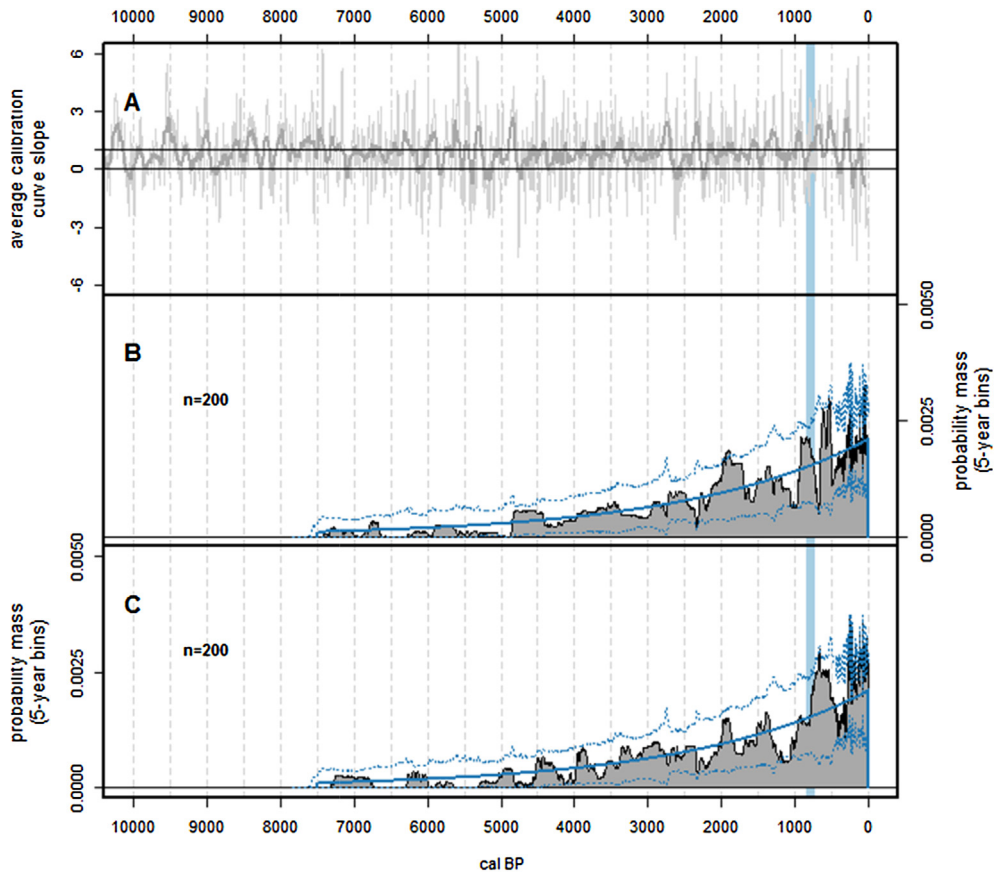

Fig. 10. *A*: 10- and 100-year moving average of the IntCal13 (Reimer et al., 2013) calibration curve slope (light and dark gray curves, respectively). *B–C*:Two MC simulations (n = 200), identically generated from a truncated exponential distribution as in Fig. 8. 2.5% and 97.5% inter-run boundaries from Fig. 8 are repeated here (dotted blue line) for comparison. Light blue vertical bar demarcates the timespan 850–750 cal BP, as in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difficult to distinguish between true demographic structures and non-demographic ones in small-sample *tfd*s, including *spd*s. These observations recommend against any attempt to address hypotheses about short-run demographic processes through ^{14}C-supported dTFA, particularly when working with modest sample sizes. Simply put, the *tfd*s that such analyses produce are too sensitive to the various DGPs that generate them to validly influence our assessment of such high-resolution hypotheses. The disputed demographic dimension of the Kachemak–Koniag transition provides a case in point. On the one hand, a scarcity of site occupations during the century 750–650 cal BP is evident in *tfd*s based on what data are currently available (Figs. 1 and 2). On the other hand, a comparison of Fig. 10:A and 10:B illustrates that even when two *spd*s are generated from an identical DGP, these will nevertheless exhibit idiosyncratic structures relative to their underlying *TFD* and disparate structures relative to each other. The first simulated *spd* (Fig. 10:B) exhibits a peak-trough-peak sequence over the interval 850 and 550 cal BP, resembling those structures observed in the Kodiak *spd* over the same interval (Fig. 2), whereas the second simulation (Fig. 10:C) instead exhibits a dramatic increase over the same interval. Conversely, the underlying *TFD* shows stable and gradual growth over the same interval, proportional to a population growing at a constant rate. In short, assuming for the sake of argument that the exponentially increasing *TFD* in Fig. 10 is representative of population dynamics in Kodiak over the period of interest, its empirical expression should nevertheless be expected to exhibit structures consistent with alternative accounts, particularly given the modest sample size available for the region. In light of the multiple stochasticities that affect small-sample *spd*s, a suspension of judgment regarding the demographic dimension of the Kachemak–Koniag transition is warranted, reprising Mills' (1994) warning of two decades ago.

Cautionary tales notwithstanding, the patterns identified in Section 3 can also be used to inform our efforts to identify effective strategies for mitigating some of the non-demographic error, both systematic and random, that troubles *spd*s. To begin with, the observation that variation decreases as sample size increases is neither surprising nor novel. In the realm of dTFA, Chamberlain (2006, 2009) and Surovell and Brantingham (2007) have already insisted that *tfd*s can only be treated as robust approximations of their underlying *TFD*s given particularly large or dense samples. Even so, our sense of what qualifies as a large or dense sample has not yet been well-calibrated. As Michczyńska and Pazdur (2004) and Williams (2012) have warned, and as the above simulations reconfirm, marked structural disparities between pairs of *spd*s identically generated from the same *TFD* can be quite dramatic for sample sizes less than 200 to 500. By extension, we should expect the story to change as sample size increases. Consequently, we should embrace every opportunity to increase sample size as a means of increasing the robustness of our *TFD* estimations, when opportunities exist to conduct further field surveys or revisit museum collections.

While the relationship between *spd* morphology and calibration curve slope has long been recognized, descriptions of this systematic error have remained incomplete until now. Michczyńska and colleagues' assertion that "the influence of the shape of the

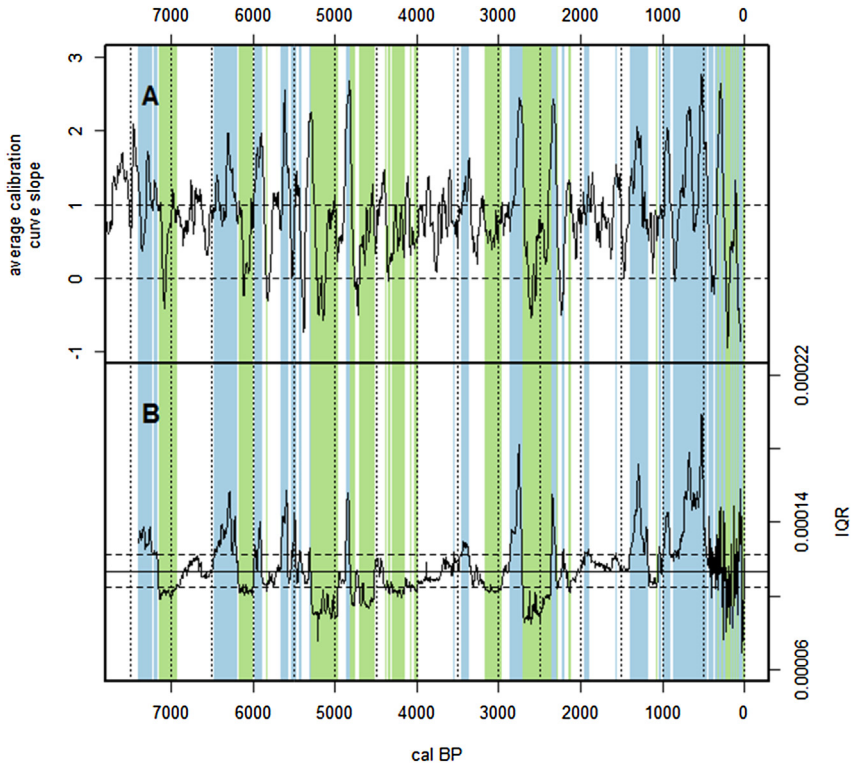

Fig. 11. *A*: 100-year moving average of the IntCal13 (Reimer et al., 2013) calibration curve slope. *B*: Inter-run interquartile ranges (IQRs) for large-sample *spd*s ($n = 2000$) based on the uniform distribution. The solid horizontal line is the median IQR value. Dashed horizontal lines bound the mid-50% range of IQRs. Light blue bars identify those temporal intervals exhibiting exceptionally high IQRs (at or above the 75% boundary), while light green bars identify those exhibiting the lowest (at or below the 25% boundary). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

calibration curve on the 95% confidence intervals is clearly visible" (2004: 739) lacks specificity, though it is powerfully illustrated in their Fig. 7. Conversely, those statements that have been more specific (see Bamforth and Grund, 2012; McFadgen et al., 1994; Williams, 2012) have also been mistaken in asserting that steep calibration curve slopes necessarily produce anomalously exaggerated peak structures in *spd*s whereas gradual slopes and plateaus dampen real peaks. We can now revise this understanding, recognizing that calibration interference manifests primarily as a time-dependent fluctuation in inter-run variance and that other stochastic processes determine whether this variance manifests as peaks or troughs. By implication, we should have greater confidence in the accuracy of *spd* morphology over those temporal intervals corresponding with low calibration curve slopes than those corresponding with high. For the range of time considered in this paper, Fig. 11:B identifies several periods that are most reliable (green bars) and those that are least (blue bars). In particular, the latter half of the 5th millennium cal BP and the five centuries centering on 2500 cal BP are especially reliable, whereas the intervals 7500–7100; 6500–6200; 5600–5400; 2900–2750; and especially 1500–500 cal BP show the greatest inter-run variance. The last 500 years BP are erratic, this being a function of noticeably low variance in the calibration curve over this terminal period (see Fig. 9). That the Kachemak–Koniag transition transpired during a period characterized by pronounced inter-run *spd* variance once again recommends that we suspend judgment regarding the demographic information that our dTFA offers for this interval.

The solution to calibration interference recommended by Williams (2012) is to smooth out *spd*s with a moving average, using a bandwidth of 500 for the interval 0–11,000 cal BP or of 800 for 11,000–50,000 cal BP (see also Ozainne et al., 2014; Peros et al., 2010; Shennan et al., 2013). This coarse-grained approach, already anticipated in McFadgen and colleagues' (1994) work, reverses the sorting of sample density that results from calibration. Of equal importance, although unrecognized by Williams, is that this smoothing method is operationally identical to kernel density estimation, a nonparametric method for estimating the probability density functions underlying samples (Baxter and Beardah, 1997; Sheather, 2004). The goal of kernel density estimation is to achieve an optimal balance between the reduction of sampling error and the preservation of structures original to the underlying probability distribution. In the future, the MC framework presented here will be used to pursue optimized algorithms for kernel density estimation in *spd*-based dTFA. The application of such 'objective' algorithms would go far in neutralizing the potential for analytical duplicity (Whallon's [1987] 'sleights of hand'). Additionally, because kernel density estimation algorithms tend to adapt to sample size and density, their application would lessen the need to pursue larger samples, though at the expense of our ability to detect fine-grained structures such as that suggested by Maschner et al. (2009b).

One approach to dTFA that has recently emerged is the measurement of proxy growth rate estimates from *tfd* curves (Crema, 2012; Collard et al., 2010; Peros et al., 2010), analogous to the measurement of population growth from $N(t)$ curves (Preston et al., 2001; Skalski et al., 2005). Such measurements however require relatively precise $N(t)$ estimates at two points along the curve, which are not easy to come by for paleopopulations. When $N(t)$ estimates for paleopopulations are imprecise, an alternative approach suggested by Cohen (1995: 77–78) is to calculate a range of plausible growth rates based on maxima and minima for the two $N(t)$ values bracketing the interval of interest, if such information is

available. Crema's (2012) estimation of the slope dynamics of a *tfd* comprising mid-Holocene Japanese housepits followed this general approach. The time-dependence of calibration interference once again suggests that if we are to derive growth rate estimates from *spd*s, the most accurate estimates will be for those intervals bracketed by ages exhibiting low inter-run variance (i.e., those designated by green bars in Fig. 11). The application of optimized KDEs should also improve the accuracy of such estimates.

Finally, while the simulations presented in this paper demonstrate that the four DGPs they model produce structures in individual *spd*s that mimic genuine demographic structures with worrisome regularity, it has been suggested elsewhere that the identification of temporally coincident structures between *tfd*s from adjacent study regions may be employed to cross-validate the demographic interpretation of such structures in each *tfd* included in the comparison (Hinz et al., 2012; Maschner et al., 2009b: 48; Maschner, personal communication). Validation and identification of constraints on this approach await future exploration.

By design, the results of this MC experiment do not explore the influence of other model system parameters on variability between runs, particularly more dispersed *TFD*s and a greater or more variable degree of instrument measurement error. As suggested by Michczyńska and Pazdur (2004) and by Williams (2012), variability in both of these parameters may warrant an adaptive approach to sample size threshold. Alternatively, an adaptive kernel density estimation algorithm could mitigate the effects of variability in these parameters. Similarly, the truncated uniform and exponential *TFD*s explored in this paper were chosen for their structural simplicity, enabling the straightforward identification of the non-demographic structures that emerge in *spd*s. Inputting more structurally complex *TFD*s into the MC simulation, which would emulate population size curves more realistically (see Bamforth and Grund, 2012; Contreras and Meadows, in press; Rhode et al., in press), would allow us to explore in greater depth the problem of equifinality that exists between real demographic structures and impostors.

5. Conclusion

The MC simulation framework presented in this paper is intended to build on the precedent of these past studies, first by automating the simulation algorithm and thereby allowing for larger and more robust simulations, and second by allowing the researcher to experimentally manipulate the many parameters that influence *tfd*s. To illustrate the power of this approach, I chose to evaluate the impact of sample size and *TFD* shape on *spd* structure, holding other parameters of the system constant. This experiment reveals a more nuanced interaction between calibration and stochastic processes than previously articulated: calibration interference expresses itself as a time-dependent fluctuation in the degree of inter-run variability. The stochastic processes, on the other hand, determine the direction and degree of such variability. Several solutions, including sample size increase, a focus on the least variable temporal intervals, and kernel density estimation, may be effective in mitigating these problems, although the last solution awaits further exploration. Future work is also warranted to explore the impacts of other parameters of the model system (chiefly instrument error and *TFD* scale and location) on the structure of ^{14}C *spd*s.

Acknowledgments

Sincere thanks go to: my colleague Stephanie Lee at CSSCR for her help in improving the efficiency of my R code; organizers of the University of Washington Department of Anthropology's 2014 S.T.A.R. conference for providing an opportunity to present early results of this research; Patrick Saltonstall, Amy Steffian, Molly Odell, and Donald Clark for providing contextual information about several ^{14}C age estimates included in Fig. 1; and Patrick Saltonstall, Ben Fitzhugh, and Marcos Llobera for their insights on this work throughout its conceptualization, development, and implementation. This paper has benefited greatly from the comments of Ben Fitzhugh, Amy Steffian, Herb Maschner, and two anonymous reviewers. All errors are solely the responsibility of the author.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jas.2014.10.013.

References

Adams, J.M., Foote, G.R., Otte, M., 2001. Could pre-Last Glacial Maximum humans have existed in North America undetected? An interregional approach to the question. Curr. Anthropol. 42 (4), 563–566. http://dx.doi.org/10.1086/322546.

Albanese, J.P., Frison, G.C., 1995. Cultural and landscape change during the Middle Holocene, Rocky Mountain area, Wyoming and Montana. In: Bettis III, E.A. (Ed.), Archaeological Geology of the Archaic Period in North America, Geological Society of America Special Paper 297, pp. 1–19. Boulder. http://dx.doi.org/10.1130/SPE297-p1.

Ammerman, A.J., Cavalli-Sforza, L.L., Wagener, D.K., 1976. Toward the estimation of population growth in Old World prehistory. In: Zubrow, E.B.W. (Ed.), Demographic Anthropology: Quantitative Approaches. University of New Mexico Press, Albuquerque, pp. 27–61.

Anderson, D.G., Meeks, S.C., Goodyear, A.C., Shane Miller, D., 2008. Southeastern data inconsistent with Paleoindian demographic reconstruction. Proc. Natl. Acad. Sci. 105 (5), E108. http://dx.doi.org/10.1073/pnas.0808964105.

Anderson, S.L., Freeburg, A.K., 2013. A high-resolution chronology for the Cape Krusenstern site complex, Northwest Alaska. Arct. Anthropol. 50 (1), 49–71.

Armit, I., Swindles, G.T., Becker, K., 2013. From dates to demography in later prehistoric Ireland? Experimental approaches to the meta-analysis of large ^{14}C data-sets. J. Archaeol. Sci. 40, 433–438. http://dx.doi.org/10.1016/j.jas.2012.08.039.

Ballenger, J.A.M., Mabry, J.B., 2011. Temporal frequency distributions of alluvium in the American Southwest: taphonomic, paleohydraulic, and demographic Implications. J. Archaeol. Sci. 38, 1314–1325. http://dx.doi.org/10.1016/j.jas.2011.01.007.

Bamforth, D.B., Grund, B., 2012. Radiocarbon calibration curves, summed probability distributions, and early Paleoindian population trends in North America. J. Archaeol. Sci. 39, 1768–1774. http://dx.doi.org/10.1016/j.jas.2012.01.017.

Bartlein, P.J., Edwards, M.E., Shafer, S.L., Barker Jr., E.D., 1995. Calibration of radiocarbon ages and the interpretation of paleoenvironmental records. Quat. Res. 44, 417–424. http://dx.doi.org/10.1006/qres.1995.1086.

Barton, L.P., Brantingham, J., Ji, D., 2007. Late Pleistocene climate change and Palaeolithic cultural evolution in northern China: implications from the Last Glacial Maximum. Dev. Quat. Sci. 9, 105–128. http://dx.doi.org/10.1016/S1571-0866(07)09009-4.

Batt, C.M., Pollard, A.M., 1996. Radiocarbon calibration and the peopling of North America. In: Orna, M.V. (Ed.), Archaeological Chemistry: Organic, Inorganic and Biochemical Analysis. American Chemical Society, Washington, D.C, pp. 415–433. http://dx.doi.org/10.1021/bk-1996-0625.ch030.

Baxter, M.J., Beardah, C.C., 1997. Some archaeological applications of kernel density estimates. J. Archaeol. Sci. 24, 347–354. http://dx.doi.org/10.1006/jasc.1996.0119.

Bever, M.R., 2006. Too little, too late? The radiocarbon chronology of Alaska and the peopling of the New World. Am. Antiq. 71 (4), 595–620. http://dx.doi.org/10.2307/40035881.

Bloom, D.E., 2011. 7 billion and counting. Science 333, 562–569. http://dx.doi.org/10.1126/science.1209290.

Brantingham, P.J., Kerry, K.W., Krivoshapkin, A.I., Kuzmin, Y.V., 2004. Time-space dynamics in the Early Upper Paleolithic of Northeast Asia. In: Madsen, D.B. (Ed.), Entering America: Northeast Asia and Beringia before the Late Glacial Maximum. University of Utah Press, Salt Lake City, pp. 255–283.

Bronk Ramsey, C., 2009a. Bayesian analysis of radiocarbon dates. Radiocarbon 51 (1), 337–360.

Bronk Ramsey, C., 2009b. Dealing with outliers and offsets in radiocarbon dating. Radiocarbon 51 (3), 1023–1045.

Buchanan, B., Collard, M., Edinborough, K., 2008. Paleoindian demography and the extraterrestrial impact hypothesis. Proc. Natl. Acad. Sci. 105 (33), 11651–11654. http://dx.doi.org/10.1073/pnas.0803762105.

Buchanan, B., Hamilton, M., Edinborough, K., O'Brien, M.J., Collard, M., 2011. A comment on Steele's (2010) "Radiocarbon dates as data: quantitative strategies for estimating colonization front speeds and event densities." J. Archaeol. Sci. 38, 2116–2122. http://dx.doi.org/10.1016/j.jas.2011.02.026.

Chamberlain, A., 2006. Demography in Archaeology. Cambridge University Press, Cambridge.

Chamberlain, A., 2009. Archaeological demography. Hum. Biol. 81 (2–3), 275–286. http://dx.doi.org/10.3378/027.081.0309.
Christian, C.P., Casella, G., 2010. Introducing Monte Carlo Methods with R. Springer, New York. http://dx.doi.org/10.1007/978-1-4419-1576-4.
Clark, D.W., 1988. Pacific Eskimo encoded precontact history. In: Shaw, R.D., Harritt, R.K., Dumond, D.E. (Eds.), The Late Prehistoric Development of Alaska's Native People, Alaska Anthropological Association Monograph Series #4, pp. 211–223. Anchorage.
Clark, D.W., 1992. "Only a skin boat load or two": the role of migration in Kodiak prehistory. Arct. Anthropol. 29 (1), 2–17.
Clark, D.W., 1998. Kodiak Island: the later cultures. Arct. Anthropol. 35 (1), 172–186.
Clevis, Q., Tucker, G.E., Lock, G., Lancaster, S.T., Gasparini, N., Desitter, A., Bras, R.L., 2006. Geoarchaeological simulation of meandering river deposits and settlement distributions: a three-dimensional approach. Geoarchaeol. Int. J. 21 (8), 843–874. http://dx.doi.org/10.1002/gea.20142.
Coale, A., 1974. The history of the human population. Sci. Am. 231 (3), 40–51. http://dx.doi.org/10.1038/scientificamerican0974-40.
Cohen, J.E., 1995. How Many People Can the Earth Support? WW Norton, New York.
Collard, M., Buchanan, B., Edinborough, K., 2008. Reply to Anderson et al., Jones, Kennett and West, Culleton, and Kennett, et al.: further evidence against the extraterrestrial impact hypothesis. Proc. Nat. Acad. Sci. 105 (50), E112–E114. http://dx.doi.org/10.1073/pnas.0810765106.
Collard, M., Buchanan, B., Hamilton, M.J., O'Brien, M.J., 2010a. Spatiotemporal dynamics of the Clovis-Folsom transition. J. Archaeol. Sci. 37, 2513–2519. http://dx.doi.org/10.1016/j.jas.2010.05.011.
Collard, M., Edinborough, K., Shennan, S., Thomas, M.G., 2010b. Radiocarbon evidence indicates that migrants introduced farming to Britain. J. Archaeol. Sci. 37, 866–870. http://dx.doi.org/10.1016/j.jas.2009.11.016.
Contreras, D.A., Meadows, J., 2014. Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. J. Archaeol. Sci. (in press) http://dx.doi.org/10.1016/j.jas.2014.05.030.
Crombé, P., Robinson, E., 2014. ^{14}C dates as demographic proxies in Neolithisation models of northwestern Europe: a critical assessment using Belgium and northeast France as a case-study. J. Archaeol. Sci. (in press) http://dx.doi.org/10.1016/j.jas.2014.02.001.
Culleton, Brendan J., 2008. Crude demographic proxy reveals nothing about Paleoindian population. Proc. Natl. Acad. Sci. 105 (50), E111. http://dx.doi.org/10.1073/pnas.0809092106.
Crema, E.R., 2012. Modelling temporal uncertainty in archaeological analysis. J. Archaeol. Method Theory 19, 440–461. http://dx.doi.org/10.1007/s10816-011-9122-3.
Dean, J.S., 1978. Independent dating in archaeological analysis. Adv. Archaeol. Method Theory 1, 223–255.
Delgado, M., Aceituno, F.J., Barrientos, G., 2014. ^{14}C data and the early colonization of Northwest South America: a critical assessment. Quat. Int. http://dx.doi.org/10.1016/j.quaint.2014.09.011 (in press).
Dumond, D.E., 1965. On Eskaleutian linguistics, archaeology, and prehistory. Am. Anthropol. 67, 1108–1115. http://dx.doi.org/10.1525/aa.1965.67.5.02a00080.
Dumond, D.E., 1987. The Eskimos and Aleuts. Thames and Hudson, London.
Dumond, D.E., 1988. The Alaska Peninsula as superhighway: a comment. In: Shaw, R.D., Harritt, R.K., Dumond, D.E. (Eds.), The Late Prehistoric Development of Alaska's Native People, Alaska Anthropological Association Monograph Series #4, pp. 379–388. Anchorage.
Dumond, D.E., 1991. The Uyak site in regional prehistory: the cultural evidence. In: Dumond, D., Scott, G.R. (Eds.), The Uyak Site on Kodiak Island: its Place in Alaskan Prehistory. University of Oregon Anthropological Papers, 44, Eugene, pp. 57–114.
Dumond, D.E., 2005. A Naknek Chronicle: Ten Thousand Years in a Land of Lakes and Rivers and Mountains of Fire. U.S. Department of the Interior, Washington, D.C. National Park Service, Katmai National Park and Preserve.
Ellison, P.T., Bogin, B., O'Rourke, M.T., 2012. Demography part 2: population growth and fertility regulation. In: Stinson, S., Bogin, B., O'Rourke, D. (Eds.), Human Biology: an Evolutionary and Biocultural Perspective, second ed. John Wiley & Sons, Inc, Hoboken, pp. 757–803. http://dx.doi.org/10.1002/9781118108062.ch15.
Erlandson, J., Crowell, A., Wooley, C., Haggarty, J., 1992. Spatial and temporal patterns in Alutiiq paleodemography. Arct. Anthropol. 29 (2), 42–62.
Erlandson, J.M., Rick, T.C., Kennett, D.J., Walker, P.L., 2001. Dates, demography, and disease: cultural contacts and possible evidence for Old World epidemics among the protohistoric Island Chumash. Pac. Coast Archaeol. Soc. Q. 37 (3), 11–26.
Esdale, J.A., 2008. A current synthesis of the Northern Archaic. Arct. Anthropol. 45 (2), 3–38. http://dx.doi.org/10.1353/arc.0.0006.
Faught, M.K., 2008. Archaeological roots of human diversity in the New World: a compilation of accurate and precise radiocarbon ages from earliest sites. Am. Antiq. 73 (4), 670–698.
Fiedel, S.J., Kuzmin, Y.V., 2007. Radiocarbon date frequency as an index of intensity of Paleolithic occupation of Siberia: did humans react predictably to climate oscillations? Radiocarbon 49 (2), 741–756.
Firestone, R.B., West, A., Kennett, J.P., Becker, L., Bunch, T.E., Revay, Z.S., Schultz, P.H., Belgya, T., Hennett, D.J., Erlandson, J.M., Dickenson, O.J., Goodyear, A.C., Harris, R.S., Howard, G.A., Kloosterman, J.B., Lechler, P., Mayewski, P.A., Montgomery, J., Poreda, R., Darrah, T., Que Hee, S.S., Smith, A.R., Stich, A., Topping, W., Wittke, J.H., Wolbach, W.S., 2007. Evidence for an extraterrestrial impact 12,900 years ago that contributed to the megafaunal extinctions and the Younger Dryas cooling. Proc. Natl. Acad. Sci. 104 (41), 16016–16021. http://dx.doi.org/10.1073/pnas.0706977104.
Fitzhugh, B., 2003. The Evolution of Complex Hunter-Gatherers: Archaeological Evidence from the North Pacific. Kluwer Academic/Plenum Publishers, New York.
Gamble, C., Davies, W., Pettitt, P., Hazelwood, L., Richards, M., 2005. The archaeological and genetic foundations of the European population during the Late Glacial: implications for 'agricultural thinking'. Camb. Archaeol. J. 15 (2), 193–223. http://dx.doi.org/10.1017/S0959774305000107.
Gamble, C., Davies, W., Pettitt, P., Richards, M., 2004. Climate change and evolving human diversity in Europe during the Last Glacial. Philos. Trans. Biol. Sci. 359 (1442), 243–254. http://dx.doi.org/10.1098/rstb.2003.1396.
Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G.K., Wilmoth, J., 2014. World population stabilization unlikely this century. Science 346, 234–237. http://dx.doi.org/10.1126/science.1257469.
Gil, A., Zárate, M., Neme, G., 2005. Mid-Holocene paleoenvironments and the archeological record of southern Mendoza, Argentina. Quat. Int. 132, 81–94. http://dx.doi.org/10.1016/j.quaint.2004.07.014.
Gkiasta, M., Russell, T., Shennan, S., Steele, J., 2003. Neolithic Transition in Europe: the radiocarbon record revisited. Antiquity 77 (295), 45–62.
González-Sampériz, P., Utrilla, P., Mazo, C., Valero-Garcés, B., Sopena, M.C., Morellón, M., Sebastián, M., Moreno, A., Martínez-Bea, M., 2009. Patterns of human occupation during the early Holocene in the Central Ebro Basin (NE Spain) in response to the 8.2 ka climatic event. Quat. Res. 71, 121–132. http://dx.doi.org/10.1016/j.yqres.2008.10.006.
Grier, C., 2006. Temporality in Northwest coast households. In: Sobel, E.A., Gahr, D.A.T., Ames, K.M. (Eds.), Household Archaeology on the Northwest Coast, International Monographs in Prehistory Archaeological Series 16, pp. 97–119. Ann Arbor.
Guilderson, T.P., Reimer, P.J., Brown, T.A., 2005. The boon and bane of radiocarbon dating. Science 307, 362–364. http://dx.doi.org/10.1126/science.1104164.
Guthrie, R.D., 2006. New carbon dates link climatic change with human colonization and Pleistocene extinctions. Nature 441, 207–209. http://dx.doi.org/10.1038/nature04604.
Haggarty, J.C., Wooley, C.B., Erlandson, J.M., Crowell, A., 1991. The 1990 Exxon Cultural Resource Program: Site Protection and Maritime Cultural Ecology in Prince William Sound and the Gulf of Alaska. Exxon Shipping Company and Exxon Company, USA, Anchorage.
Hassan, F.A., 1981. Demographic Archaeology. Academic Press, New York.
Haynes Jr., C.V., 1969. The earliest Americans. Science 166, 709–715. http://dx.doi.org/10.1126/science.166.3906.709.
Haynes, G., Anderson, D.G., Ferring, C.R., Fiedel, S.J., Grayson, D.K., Haynes Jr., C.V., Holliday, V.T., Huckell, B.B., Kornfeld, M., Meltzer, D.J., Morrow, J., Surovell, T., Waguespack, N.M., Wigand, P., Yohe II, R.M., 2007. Comment on "Redefining the age of Clovis: implications for the peopling of the Americas." Science 317, 320b. http://dx.doi.org/10.1126/science.1141960.
Hinz, M., Feeser, I., Sjögren, K.-G., Müller, J., 2012. Demography and the intensity of cultural activities: an evaluation of Funner Beaker Societies (4200–2800 cal BC). J. Archaeol. Sci. 39, 3331–3340. http://dx.doi.org/10.1016/j.jas.2012.05.028.
Hirschman, C., 2005. Population and society: historical trends and future prospects. In: Calhoun, C., Rojek, C., Turner, B.S. (Eds.), The Sage Handbook of Sociology. Sage Publications, London, pp. 381–402. http://dx.doi.org/10.4135/9781848608115.n23.
Holdaway, S., Fanning, P., Rhodes, E., 2008. Challenging intensification: human-environment interactions in the Holocene geoarchaeological record from western New South Wales, Australia. Holocene 18 (3), 403–412. http://dx.doi.org/10.1177/0959683607087930.
Holdaway, S., Porch, N., 1995. Cyclical patterns in the Pleistocene human occupation of Southwest Tasmania. Archaeol. Ocean. 30 (2), 74–82.
Housley, R.A., Gamble, C.S., Street, M., Pettitt, P., 1997. Radiocarbon evidence for the Lateglacial human recolonisation of northern Europe. Proc. Prehist. Soc. 63, 25–54.
Hrdlička, A., 1944. The Anthropology of Kodiak Island. Press of the Wistar Institute of Anatomy and Biology, Philadelphia.
Hutchison, I., McMillan, A.D., 1997. Archaeological evidence for village abandonment associated with late Holocene earthquakes at the northern Cascadia subduction zone. Quat. Res. 48, 79–87. http://dx.doi.org/10.1006/qres.1997.1890.
Jones, T.L., 2008. California archaeological record consistent with Younger Dryas disruptive event. Proc. Natl. Acad. Sci. 105 (50), E109. http://dx.doi.org/10.1073/pnas.0808976106.
Jordan, R.H., Knecht, R.A., 1988. Archaeological research on western Kodiak Island, Alaska: the development of Koniag culture. In: Shaw, R.D., Harritt, R.K., Dumond, D.E. (Eds.), The Late Prehistoric Development of Alaska's Native People, Alaska Anthropological Association Monograph Series #4, pp. 225–306. Anchorage.
Kelly, R.L., Surovell, T.A., Shuman, B.N., Smith, G.M., 2013. A continuous climatic impact on Holocene human population in the Rocky Mountains. Proc. Natl. Acad. Sci. 110 (2), 443–447. http://dx.doi.org/10.1073/pnas.1201341110.
Kennett, D.J., Stafford Jr., T.W., Southon, J., 2008. Standards of evidence and Paleoindian demographics. Proc. Natl. Acad. Sci. 105 (50), E107. http://dx.doi.org/10.1073/pnas.0808960106.
Kennett, J.P., West, A., 2008. Biostratigraphic evidence supports Paleoindian population disruption at ≈12.9 ka. Proc. Natl. Acad. Sci. 105 (50), E110. http://dx.doi.org/10.1073/pnas.0809004106.

Kerr, T.R., McCormick, F., 2014. Statistics, sunspots and settlement: influences on sum of probability curves. J. Archaeol. Sci. 41, 493–501. http://dx.doi.org/10.1016/j.jas.2013.09.002.
Kim, J.C., Bae, C.J., 2010. Radiocarbon dates documenting the Neolithic-Bronze age transition in Korea. Radiocarbon 52 (2), 483–492.
Kirch, P.V., 1985. Feathered Gods and Fishhooks: an Introduction to Hawaiian Archaeology and Prehistory. University of Hawai'i Press, Honolulu.
Kirch, P.V., Asner, G., Chadwick, O.A., Field, J., Ladefoged, T., Lee, C., Puleston, C., Tuljapurkar, S., Vitousek, P.M., 2012. Building and testing models of long-term agricultural intensifications and population dynamics: a case study from the Leeward Kohala Field System, Hawai'i. Ecol. Model. 227, 18–28. http://dx.doi.org/10.1016/j.ecolmodel.2011.11.032.
Knecht, R.A., 1995. The Late Prehistory of the Alutiiq People: Culture Change on the Kodiak Archipelago from 1200–1750 A.D (Ph.D. dissertation).
Kuzmin, Y.V., Keates, S.G., 2005. Dates are not just data: Paleolithic settlement patterns in Siberia derived from radiocarbon records. Am. Antiq. 70 (4), 773–789. http://dx.doi.org/10.2307/40035874.
Lee, C.T., Puleston, C.O., Tuljapurkar, S., 2009. Population and prehistory III: food-dependent demography in variable environments. Theor. Popul. Biol. 76, 179–188. http://dx.doi.org/10.1016/j.tpb.2009.06.003.
Lee, C.T., Tuljapurkar, S., 2008. Population and prehistory I: food-dependent population growth in constant environments. Theor. Popul. Biol. 73, 473–482. http://dx.doi.org/10.1016/j.tpb.2008.03.001.
Lepofsky, D., Lertzman, K., Hallett, D., Mathewes, R., 2005. Climate change and culture change on the southern coast of British Columbia 2400–1200 Cal. B.P.: an hypothesis. Am. Antiq. 70 (2), 267–293. http://dx.doi.org/10.2307/40035704.
Liedgren, L.G., Bergman, I.M., Hörnberg, G., Zackrisson, O., Hellberg, E., Östland, L., DeLuca, T.H., 2007. Radiocarbon dating of prehistoric hearths in alpine northern Sweden: problems and possibilities. J. Archaeol. Sci. 34, 1276–1288. http://dx.doi.org/10.1016/j.jas.2006.10.018.
Louderback, L., Grayson, D.K., Llobera, M., 2011. Middle-Holocene climates and human population densities in the Great Basin, western USA. Holocene 21 (2), 366–373. http://dx.doi.org/10.1177/0959683610374888.
McFadgen, B.G., Knox, F.B., Cole, T.R.L., 1994. Radiocarbon calibration curve variations and their implications for the interpretation of New Zealand prehistory. Radiocarbon 36 (2), 221–236.
MacPhee, R.D.E., Tikhonov, A.N., Mol, D., de Marliave, C., van der Plicht, H., Greenwood, A.D., Flemming, C., Agenbroad, L., 2002. Radiocarbon chronologies and extinction dynamics of the Late Quaternary mammalian megafauna of the Taimyr Peninsula, Russian Federation. J. Archaeol. Sci. 29, 1017–1042. http://dx.doi.org/10.1006/jasc.2001.0802.
Mandel, R.D., 1995. Geomorphic controls of the Archaic record in the Central Plains of the United States. In: Bettis III, E.A. (Ed.), Archaeological Geology of the Archaic Period in North America, Geological Society of America Special Paper 297, pp. 37–66. Boulder. http://dx.doi.org/10.1130/SPE297-p37.
Martínez, P.V.C., Suriñach, S.G., Marcén, P.G., Lull, V., Pérez, R.M., Herrada, C.R., 1997. Radiocarbon dating and the prehistory of the Balearic Islands. Proc. Prehist. Soc. 63, 55–86.
Marwick, B., 2009. Change or decay? An interpretation of late Holocene archaeological evidence from the Hamersley Plateau, western Australia. Archaeol. Ocean. 44 (Suppl.), 16–22. http://dx.doi.org/10.1002/j.1834-4453.2009.tb00064.x.
Maschner, H.D.G., Betts, M.W., Cornell, J., Dunne, J.A., Finney, B., Huntly, N., Jordan, J.W., King, A.A., Misarti, N., Reedy-Maschner, K.L., Russell, R., Tews, A., Wood, S.A., Benson, B., 2009a. An introduction to biocomplexity of Sanak Island, western Gulf of Alaska. Pac. Sci. 63 (4), 673–709. http://dx.doi.org/10.2984/049.063.0410.
Maschner, H., Finney, B., Jordan, J., Misarti, N., Tews, A., Knudsen, B., 2009b. Did the North Pacific ecosystem collapse in AD 1250? In: Maschner, H., Mason, O., McGhee, R. (Eds.), The Northern World AD 900-1400. The University of Utah Press, Salt Lake City, pp. 33–57.
Mason, O.K., Bowers, P.M., Hopkins, D.M., 2001. The Early Holocene Milankovitch thermal maximum and humans: adverse conditions for the Denali complex of eastern Beringia. Quat. Sci. Rev. 20, 525–548. http://dx.doi.org/10.1016/S0277-3791(00)00100-1.
Mellars, P., French, J.C., 2010. Tenfold population increase in western Europe at the Neanderthal-to-modern human transition. Science 333, 623–627. http://dx.doi.org/10.1126/science.1206930.
Michczyńska, D.J., Michczyński, A., Pazdur, A., Żurek, S., 2003. ^{14}C dates of peat for reconstruction of environmental changes in the past. Geochronometria 22, 47–54.
Michczyńska, D.J., Pazdur, A., 2004. Shape analysis of cumulative probability density function of radiocarbon dates set in the study of climate change in Late Glacial and Holocene. Radiocarbon 46 (2), 733–744.
Micklin, M., Poston Jr., D.L., 2005. Prologue: the demographer's ken: 50 years of growth and change. In: Poston, D.L., Micklin, M. (Eds.), Handbook of Population. Springer, New York, pp. 1–15.
Miller, G.H., Magee, J.W., Johnson, B.J., Fogel, M.L., Spooner, N.A., McCulloch, M.T., Ayliffe, L.K., 1999. Pleistocene extinction of *Genyornis newtoni*: human impact on Australian megafauna. Science 283, 205–208. http://dx.doi.org/10.1126/science.283.5399.205.
Mills, R.O., 1994. Radiocarbon calibration of archaeological dates from the central Gulf of Alaska. Arct. Anthropol. 31 (1), 126–149.
Moreno, A., Santoro, C.M., Latore, C., 2009. Climate change and human occupation in the northernmost Chilean Altiplano over the last ca. 11500 cal. a BP. J. Quat. Sci. 24 (4), 373–382. http://dx.doi.org/10.1002/jqs.1240.
Mullen, P.O., 2012. An archaeological test of the effects of the White River Ash eruptions. Arct. Anthropol. 49 (1), 35–44. http://dx.doi.org/10.1353/arc.2012.0013.
Mulrooney, M.A., 2013. An island-wide assessment of the chronology of settlement and land use on Rapa Nui (Easter Island) based on radiocarbon data. J. Archaeol. Sci. 40, 4377–4399. http://dx.doi.org/10.1016/j.jas.2013.06.020.
Munoz, S.E., Gajewski, K., Peros, M.C., 2010. Synchronous environmental and cultural change in the prehistory of the northeastern United States. Proc. Natl. Acad. Sci. 107 (51), 22008–22013. http://dx.doi.org/10.1073/pnas.1005764107.
Naudinot, N., Tomasso, A., Tozzi, C., Peresani, M., 2014. Changes in mobility patterns as a factor of ^{14}C date density variation in the Late Epigravettian of Northern Italy and Southeastern France. J. Archaeol. Sci. (in press) http://dx.doi.org/10.1016/j.jas.2014.05.021.
Newell, C., 1988. Methods and Models in Demography. The Guilford Press, New York.
Ozainne, S., Lespez, L., Garnier, A., Ballouche, A., Neumann, K., Pays, O., Huysecom, E., 2014. A question of timing: spatio-temporal structure and mechanisms of early agriculture expansion in West Africa. J. Archaeol. Sci. 50, 359–368. http://dx.doi.org/10.1016/j.jas.2014.07.025.
Pazdur, M.F., Michczyńska, D.J., 1989. Improvement of the procedure for probabilistic calibration of radiocarbon. Radiocarbon 31 (3), 824–832.
Peros, M.C., Munoz, S.E., Gajewski, K., Viau, A.E., 2010. Prehistoric demography of North America inferred from radiocarbon data. J. Archaeol. Sci. 37, 656–664. http://dx.doi.org/10.1016/j.jas.2009.10.029.
Pettitt, P.B., Davies, W., Gamble, C.S., Richards, M.B., 2003. Palaeolithic radiocarbon chronology: quantifying our confidence beyond two half-lives. J. Archaeol. Sci. 30, 1685–1693. http://dx.doi.org/10.1016/S0305-4403(03. http://dx.doi.org/10.1016/S0305-4403(03)00070-0.
Potter, B.A., 2008. Radiocarbon chronology of central Alaska: technological continuity and economic change. Radiocarbon 50 (2), 181–204.
Preston, S.H., Heuveline, P., Guillot, M., 2001. Demography: Measuring and Modeling Population Processes. Blackwell Publishing, Malden, MA.
Puleston, C.O., Tuljapurkar, S., 2008. Population and prehistory II: space-limited human populations in constant environments. Theor. Popul. Biol. 74, 147–160. http://dx.doi.org/10.1016/j.tpb.2008.05.007.
Rasic, J.T., Matheus, P.E., 2007. A reconsideration of purported Holocene bison bones from northern Alaska. Arctic 50 (4), 381–388. http://dx.doi.org/10.14430/arctic195.
Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T.J., Hoffmann, D.L., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Niu, M., Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Staff, R.A., Turney, C.S.M., van der Plicht, J., 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. Radiocarbon 55 (4), 1869–1887.
Rhode, D., Brantingham, P.J., Perreault, C., Madsen, D.B., 2014. Mind the gaps: testing for hiatuses in regional radiocarbon date sequences. J. Archaeol. Sci. (in press) http://dx.doi.org/10.1016/j.jas.2014.02.022.
Rick, J.W., 1987. Dates as data: an examination of the Peruvian Preceramic radiocarbon record. Am. Antiq. 52 (1), 55–73. http://dx.doi.org/10.2307/281060.
Riede, F., 2008. The Laacher See-eruption (12,920 BP) and material culture change at the end of the Allerød in Northern Europe. J. Archaeol. Sci. 35, 591–599. http://dx.doi.org/10.1016/j.jas.2007.05.007.
Riede, F., 2009. Climate and demography in early prehistory: using calibrated ^{14}C dates as population proxies. Hum. Biol. 81 (2–3), 309–337. http://dx.doi.org/10.3378/027.081.0311.
Rieth, T.M., Hunt, T.L., 2008. A radiocarbon chronology for Sāmoan prehistory. J. Archaeol. Sci. 35, 1901–1927. http://dx.doi.org/10.1016/j.jas.2007.12.001.
Roberts, C.P., Casella, G., 2004. Monte Carlo Statistical Methods. Spring, New York.
Rogers, A.R., 1992. Resources and population dynamics. In: Smith, E.A., Winterhalder, B. (Eds.), Evolutionary Ecology and Human Behavior. Aldine-Transaction, New Brunswick, pp. 375–402.
Rubinstein, R.Y., Kroese, D.P., 2008. Simulation and the Monte Carlo Method. John Wiley & Sons, Hoboken.
Saltonstall, P.G., Steffian, A.F., 2006. The Archeology of Horseshoe Cove: Excavation at KOD-415, Uganik Island, Kodiak Archipelago, Alaska. Occasional Papers in Alaskan Field Archeology Number 1. Bureau of Indian Affairs Alaska Region, Office of Regional Archeology U.S. Department of the Interior.
Schiffer, M.B., 1987. Formation Processes of the Archaeological Record. University of New Mexico Press, Albuquerque.
Schmidt, I., Bradtmöller, M., Kehl, M., Pastoors, A., Tafelmaier, Y., Weninger, B., Weninger, G.-C., 2012. Rapid climate change and variability of settlement patterns in Iberia during the Late Pleistocene. Quat. Int. 274, 179–204. http://dx.doi.org/10.1016/j.quaint.2012.01.018.
Sheather, S.J., 2004. Density estimation. Stat. Sci. 19 (4), 588–597. http://dx.doi.org/10.1214/088342304000000297.
Shennan, S., Downey, S.S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., Thomas, M.G., 2013. Regional population collapse followed initial agricultural booms in Mid-Holocene Europe. Nat. Commun. 4. http://dx.doi.org/10.1038/ncomms3486.

Shennan, S., Edinborough, K., 2007. Prehistoric population history: from the Late Glacial to the Late Neolithic in central and northern Europe. J. Archaeol. Sci. 34, 1339–1345. http://dx.doi.org/10.1016/j.jas.2006.10.031.
Shott, M.J., 1992. Radiocarbon dating as a probabilistic technique: the Childers site and Late Woodland occupation in the Ohio Valley. Am. Antiq. 57 (2), 202–230. http://dx.doi.org/10.2307/280728.
Skalski, J.R., Ryding, K.E., Millspaugh, J.J., 2005. Wildlife Demography: Analysis of Sex, Age, and Count Data. Elsevier Academic Press, Amsterdam.
Smith, M.A., Ross, J., 2008. What happened at 1500–1000 cal. BP in Central Australia? Timing, impact and archaeological signatures. Holocene 18 (3), 379–388. http://dx.doi.org/10.1177/0959683607087928.
Smith, M.A., Williams, A.N., Turney, C.S.M., Cupper, M.L., 2008. Human-environment interactions in Australian drylands: exploratory time-series analysis of archaeological records. Holocene 18 (3), 389–401. http://dx.doi.org/10.1177/0959683607087929.
Sonett, C.P., Finney, S.A., 1990. The spectrum of radiocarbon. Philosophical Trans. R. Soc. Lond. A 330, 413–426.
Steele, J., 2010. Radiocarbon dates as data: quantitative strategies for estimating colonization front speeds and event densities. J. Archaeol. Sci. 37, 2017–2030. http://dx.doi.org/10.1016/j.jas.2010.03.007.
Steier, P., Rom, W., Puchegger, S., 2001. New methods and critical aspects in Bayesian mathematics for ^{14}C calibration. Radiocarbon 43 (2A), 373–380.
Story, D.A., Valastro Jr., S., 1977. Radiocarbon dating and the George C Davis site, Texas. J. Field Archaeol. 4 (1), 63–89. http://dx.doi.org/10.2307/529725.
Stuiver, M., Reimer, P.J., 1989. Histograms obtained from computerized radiocarbon age calibration. Radiocarbon 31 (3), 817–823.
Stuiver, M., Reimer, P.J., 1993. Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program. Radiocarbon 35 (1), 215–230.
Surovell, T.A., Brantingham, P.J., 2007. A note on the use of temporal frequency distributions in studies of prehistoric demography. J. Archaeol. Sci. 34, 1868–1877. http://dx.doi.org/10.1016/j.jas.2007.01.003.
Surovell, T.A., Finley, J.B., Smith, G.M., Brantingham, P.J., Kelly, R., 2009. Correcting temporal frequency distributions for taphonomic bias. J. Archaeol. Sci. 36, 1715–1724. http://dx.doi.org/10.1016/j.jas.2009.03.029.
Surovell, T., Waguespack, N., Brantingham, P.J., 2005. Global archaeological evidence for proboscidean overkill. Proc. Natl. Acad. Sci. 102 (17), 6231–6236. http://dx.doi.org/10.1073/pnas.0501947102.
Tallavaara, M., Pesonen, P., Oinonen, M., 2010. Prehistoric population history in eastern Fennoscandia. J. Archaeol. Sci. 37, 251–260. http://dx.doi.org/10.1016/j.jas.2009.09.035.
Tallavaara, M., Seppä, H., 2012. Did the mid-Holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern Fennoscandia? Holocene 22 (2), 215–225. http://dx.doi.org/10.1177/0959683611414937.
Taylor, A.K., Stein, J.K., Jolivette, S.A.E., 2011. Big sites, small sites, and coastal settlement patterns in the San Juan Islands, Washington, USA. J. Isl. Coast. Archaeol. 6, 287–313. http://dx.doi.org/10.1080/15564894.2010.504614.
Thomopoulos, N.T., 2013. Essentials of Monte Carlo Simulation: Statistical Methods for Building Simulation Models. Springer, New York.
Timpson, A., Colledge, S., Crema, E., Ediborough, K., Kerig, T., Manning, K., Thomas, M.G., Shennan, S., 2014. Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. J. Archaeol. Sci. (in press) http://dx.doi.org/10.1016/j.jas.2014.08.011.
Turney, C.S.M., Baillie, M., Palmer, J., Brown, D., 2006. Holocene climate change and past Irish societal response. J. Archaeol. Sci. 33, 34–38. http://dx.doi.org/10.1016/j.jas.2005.05.014.
Turney, C.S.M., Hobbs, D., 2006. ENSO influence on Holocene Aboriginal populations in Queensland, Australia. J. Archaeol. Sci. 33, 1744–1748. http://dx.doi.org/10.1016/j.jas.2006.03.007.
Walker, M., 2005. Quaternary Dating Methods. John Wiley & Sons, Ltd, West Sussex.
Wang, C., Lu, H., Zhang, J., Gu, Z., He, K., 2014. Prehistoric demographic fluctuations in China inferred from radiocarbon data and their linkage with climate change over the past 50,000 years. Quat. Sci. Rev. 98, 45–59. http://dx.doi.org/10.1016/j.quascirev.2014.05.015.
Ward, G.K., Wilson, S.R., 1978. Procedures for comparing and combining radiocarbon age determinations: a critique. Archaeometry 20 (1), 19–31. http://dx.doi.org/10.1111/j.1475-4754.1978.tb00208.x.
Waters, M.R., Kuehn, D.D., 1996. The geoarchaeology of place: the effect of geological processes on the preservation and interpretation of the archaeological record. Am. Antiq. 61 (3), 483–497. http://dx.doi.org/10.2307/281836.
Waters, M.R., Stafford Jr., T.W., 2007. Redefining the age of Clovis: implications for the peopling of the Americas. Science 315, 1122–1126. http://dx.doi.org/10.1126/science.1137166.
Weninger, B., 1986. High-precision calibration of archaeological radiocarbon dates. Acta Interdiscip. Archaeol. 4, 11–53.
Weninger, B., Edinborough, K., Clare, L., Jöris, O., 2011. Concepts of probability in radiocarbon analysis. Doc. Praehist. 38, 1–20. http://dx.doi.org/10.4312/dp.38.2.
West, C.F., 2011. A revised radiocarbon sequence for Karluk-1 and the implications for Kodiak Island prehistory. Arct. Anthropol. 48 (1), 80–92. http://dx.doi.org/10.1353/arc.2011.0111.
Whallon, R., 1987. Simple statistics. In: Aldenderfer, M.S. (Ed.), Quantitative Research in Archaeology: Progress and Prospects. Sage, Newbury Park, NJ, pp. 135–150.
Whitehouse, N.J., Schulting, R.J., McClatchie, M., Barratt, P., McLaughlin, T.R., Bogaard, A., Colledge, S., Marchant, R., Gaffrey, J., Bunting, M.J., 2014. Neolithic agriculture on the European western frontier: the boom and bust of early farming in Ireland. J. Archaeol. Sci. 51, 181–205. http://dx.doi.org/10.1016/j.jas.2013.08.009.
Wicks, K., Mithen, S., 2014. The impact of the abrupt 8.2 ka cold event on the Mesolithic population of western Scotland: a Bayesian chronological analysis using 'activity events' as a population proxy. J. Archaeol. Sci. 45, 240–269. http://dx.doi.org/10.1016/j.jas.2014.02.003.
Wicks, K., Pirie, A., Mithen, S.J., 2014. Settlement patterns in the late Mesolithic of western Scotland: the implications of Bayesian analysis of radiocarbon dates and inter-site technological comparisons. J. Archaeol. Sci. 41, 406–422. http://dx.doi.org/10.1016/j.jas.2013.07.003.
Williams, A.N., 2012. The use of summed radiocarbon probability distributions in archaeology: a review of methods. J. Archaeol. Sci. 39, 578–589. http://dx.doi.org/10.1016/j.jas.2011.07.014.
Williams, A., Santoro, C.M., Smith, M.A., Latorre, C., 2008. The impact of ENSO in the Atacama Desert and Australian arid zone: exploratory time-series analysis of archaeological records. Chungara, Rev. Antropol. Chil. 40, 245–259. http://dx.doi.org/10.4067/S0717-73562008000300003.
Williams, A.N., Ulm, S., Cook, A.R., Langley, M.C., Collard, M., 2013. Human refugia in Australia during the Last Glacial Maximum and Terminal Pleistocene: a geospatial analysis of the 25–12 ka Australian archaeological record. J. Archaeol. Sci. 40, 4612–4625. http://dx.doi.org/10.1016/j.jas.2013.06.015.
Williams, A.N., Ulm, S., Goodwin, I.D., Smith, M., 2010. Hunter-gatherer response to late Holocene climatic variability in northern and central Australia. J. Quat. Sci. 25 (6), 831–838. http://dx.doi.org/10.1002/jqs.1416.
Woodbridge, J., Fyfe, R.M., Roberts, N., Downey, S., Ediborough, K., Shennan, S., 2014. The impact of the Neolithic agricultural transition in Britain: a comparison of pollen-based land-cover and archaeological ^{14}C date-inferred population change. J. Archaeol. Sci. 51, 216–224. http://dx.doi.org/10.1016/j.jas.2012.10.025.